

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 12:24:58 ; Search time 29.94 seconds
(without alignments)
63.068 Million cell updates/sec

Title: US-09-493-795A-1

Perfect score: 49

Sequence: 1 xxxxxccxxxccxxxcc 17

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Database :

A_Geneseq_032802.*
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | ID | Description |
|------------|-------|-------------|----|-------------|
| 1 | 49 | 100.0 | 17 | 21 AAB21369 |
| 2 | 49 | 100.0 | 17 | 21 AAB21379 |
| 3 | 49 | 100.0 | 25 | 22 ABB39831 |
| 4 | 49 | 100.0 | 25 | 22 ABB24435 |
| 5 | 49 | 100.0 | 25 | 22 AAM50559 |
| 6 | 49 | 100.0 | 25 | 22 AAM73218 |
| 7 | 49 | 100.0 | 25 | 22 AAG4879 |
| 8 | 49 | 100.0 | 25 | 22 AAG4882 |
| 9 | 49 | 100.0 | 25 | 22 AAG4885 |
| 10 | 49 | 100.0 | 25 | 22 AAG4888 |
| 11 | 49 | 100.0 | 25 | 22 AAG4889 |

| | | | | | | |
|----|----|-------|----|----|----------|--------------------|
| 12 | 49 | 100.0 | 25 | 22 | AAG64890 | Human protein rela |
| 13 | 49 | 100.0 | 25 | 22 | AAG64891 | Human protein rela |
| 14 | 49 | 100.0 | 25 | 22 | AA19888 | Peptide #6322 enco |
| 15 | 49 | 100.0 | 25 | 22 | AA19888 | Peptide #7461 enco |
| 16 | 49 | 100.0 | 26 | 20 | AA19888 | Fragment of human |
| 17 | 49 | 100.0 | 26 | 22 | AA19888 | Cone snail O-supe |
| 18 | 49 | 100.0 | 27 | 22 | AA19888 | Cone snail O-supe |
| 19 | 49 | 100.0 | 28 | 22 | AA19888 | Novel human diagno |
| 20 | 49 | 100.0 | 29 | 15 | AA19888 | Spider venom calci |
| 21 | 49 | 100.0 | 29 | 19 | AA19888 | Analgesic peptide |
| 22 | 49 | 100.0 | 29 | 19 | AA19888 | C. gloriamaris del |
| 23 | 49 | 100.0 | 29 | 21 | AA19888 | Analgesic peptide |
| 24 | 49 | 100.0 | 29 | 21 | AA19888 | Conus gloriamaris |
| 25 | 49 | 100.0 | 29 | 22 | AA19888 | Cone snail O-supe |
| 26 | 49 | 100.0 | 29 | 22 | AA19888 | Cone snail O-supe |
| 27 | 49 | 100.0 | 29 | 22 | AA19888 | Cone snail O-supe |
| 28 | 49 | 100.0 | 30 | 15 | AA19888 | Spider venom calci |
| 29 | 49 | 100.0 | 30 | 15 | AA19888 | Spider venom calci |
| 30 | 49 | 100.0 | 30 | 15 | AA19888 | Outward K channel |
| 31 | 49 | 100.0 | 30 | 22 | AA19888 | Human polypeptide |
| 32 | 49 | 100.0 | 30 | 22 | AA19888 | Cone snail O-supe |
| 33 | 49 | 100.0 | 31 | 15 | AA19888 | Spider venom calci |
| 34 | 49 | 100.0 | 31 | 15 | AA19888 | Spider venom calci |
| 35 | 49 | 100.0 | 31 | 15 | AA19888 | Outward K channel |
| 36 | 49 | 100.0 | 31 | 19 | AA19888 | Analgesic peptide |
| 37 | 49 | 100.0 | 31 | 19 | AA19888 | Grammostola spatul |
| 38 | 49 | 100.0 | 31 | 20 | AA19888 | Chilean pink taran |
| 39 | 49 | 100.0 | 31 | 20 | AA19888 | Analgesic peptide |
| 40 | 49 | 100.0 | 31 | 20 | AA19888 | muO-conopeptide G2 |
| 41 | 49 | 100.0 | 32 | 21 | AA19888 | Cone snail alpha-c |
| 42 | 49 | 100.0 | 32 | 21 | AA19888 | Cone snail alpha-c |
| 43 | 49 | 100.0 | 35 | 21 | AA19888 | Cone snail alpha-c |
| 44 | 49 | 100.0 | 35 | 21 | AA19888 | Cone snail alpha-c |
| 45 | 49 | 100.0 | 35 | 21 | AA19888 | Cone snail alpha-c |
| 46 | 49 | 100.0 | 35 | 21 | AA19888 | Cone snail alpha-c |
| 47 | 49 | 100.0 | 35 | 21 | AA19888 | Cone snail alpha-c |
| 48 | 49 | 100.0 | 35 | 21 | AA19888 | Cone snail alpha-c |
| 49 | 49 | 100.0 | 35 | 21 | AA19888 | Cone snail alpha-c |
| 50 | 49 | 100.0 | 35 | 22 | ABG29284 | Novel human diagno |

ALIGNMENTS

RESULT 1

AAB21369

ID AAB21369 standard; Peptide: 17 AA.

XX AC AAB21369;

XX AC AAB21369;

DT 19-JAN-2001 (first entry)

XX

DE Cone snail alpha-conotoxin generic sequence SEQ ID NO: 1.

XX Cone snail; alpha-conotoxin; venom: disulphide bond; mood disorder;

KW neuronal nicotinic acetylcholine receptor; cardiovascular disorder;

KW gastric motility disorder; urinary incontinence; nicotine addiction;

XX small cell lung carcinoma.

OS Conus sp.

XX

FH Key Location/Qualifiers

FT Misc-difference 1 /label= Ile, Leu, Val, OTHER

FT /note= "des-Xaa"

FT Misc-difference 2

FT /label= Ala, Gly, OTHER

FT /note= "des-Xaa"

FT Misc-difference 3

FT /label= Gly, Trp, OTHER

FT /note= "OTHER= des-Xaa, neo-Trp, halo-Trp or another

FT unnatural aromatic amino acid. Trp may be D- or L-form"

FT Misc-difference 4

FT

The invention relates to a single exon nucleic acid probe for measuring human gene expression in a sample derived from human foetal liver. The single exon nucleic acid probes may be used for predicting, measuring and displaying gene expression in samples derived from human foetal liver. The present sequences a peptide encoded by a single exon

XX 05-JUN-2001.
XX 12-SEP-2000; 2000JP-0276083.
XX 16-SEP-1999; 99JP-0262228.
XX (TAKE) TAKEDA CHEM IND LTD.
XX WPI; 2001-499381/55.
XX N-PSDB; AAH73470.

PT New polypeptide and its DNA, useful for diagnosis, treatment and
PT prevention of infection, septicemia, drug intoxication, cancer, liver
PT and immune dysfunction -

XX Claim 10; Page 33; 41pp; Japanese.

XX The present invention provides the protein shown in SEQ ID NO: 3 in the
XX specification. It can be used for the diagnosis, treatment and prevention
XX of opportunistic infections, septicemia, drug intoxication, phthisis,
XX cancers, liver dysfunction, immune dysfunction and dyschormonism. The
XX present sequence is a peptide described in the exemplification of the
XX invention.

XX SQ Sequence 25 AA;

Query Match 100.0%; Score 49; DB 22; Length 25;
Best Local Similarity 23.5%; Pred. NO. 4.8e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

OY 1 XXXXXCXXXXXXCX 17
DB 8 ifccgchrskcmck 24

RESULT 8
AAG64882
ID AAG64882 standard; peptide; 25 AA.

XX AC AAG64882;

XX 15-OCT-2001 (first entry)

XX Murine peptide SEQ ID NO: 13.

XX Mouse; infection; septicemia; drug intoxication; phthisis; cancer;
XX liver dysfunction; immune dysfunction; dyschormonism; antibacterial.

XX Mus sp.

XX JP2001149083-A.

XX 05-JUN-2001.

XX 12-SEP-2000; 2000JP-0276083.

XX 16-SEP-1999; 99JP-0262228.

XX (TAKE) TAKEDA CHEM IND LTD.

XX WPI; 2001-499381/55.

XX N-PSDB; AAH73473.

XX New polypeptide and its DNA, useful for diagnosis, treatment and
XX prevention of infection, septicemia, drug intoxication, cancer, liver
XX and immune dysfunction -

XX Claim 3; Page 34-35; 41pp; Japanese.

XX The present invention provides the protein shown in SEQ ID NO: 3 in the
XX specification. It can be used for the diagnosis, treatment and prevention

CC of opportunistic infections, septicemia, drug intoxication, phthisis,
CC cancers, liver dysfunction, immune dysfunction and dyschormonism. The
CC present sequence is a peptide described in the exemplification of the
CC invention.

XX SQ Sequence 25 AA;

Query Match 100.0%; Score 49; DB 22; Length 25;
Best Local Similarity 23.5%; Pred. NO. 4.8e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

OY 1 XXXXXCXXXXXXCX 17
DB 8 ifccckcnnsqcgicck 24

RESULT 9
AAG64885
ID AAG64885 standard; peptide; 25 AA.

XX AC AAG64885;

XX 15-OCT-2001 (first entry)

XX Murine peptide SEQ ID NO: 18.

XX Mouse; infection; septicemia; drug intoxication; phthisis; cancer;
XX liver dysfunction; immune dysfunction; dyschormonism; antibacterial.

XX Mus sp.

XX JP2001149083-A.

XX 05-JUN-2001.

XX 12-SEP-2000; 2000JP-0276083.

XX 16-SEP-1999; 99JP-0262228.

XX (TAKE) TAKEDA CHEM IND LTD.

XX WPI; 2001-499381/55.

XX N-PSDB; AAH73476.

XX New polypeptide and its DNA, useful for diagnosis, treatment and
XX prevention of infection, septicemia, drug intoxication, cancer, liver
XX and immune dysfunction -

XX Example 4; Page 36; 41pp; Japanese.

XX The present invention provides the protein shown in SEQ ID NO: 3 in the
XX specification. It can be used for the diagnosis, treatment and prevention
XX of opportunistic infections, septicemia, drug intoxication, phthisis,
XX cancers, liver dysfunction, immune dysfunction and dyschormonism. The
XX present sequence is a peptide described in the exemplification of the
XX invention.

XX SQ Sequence 25 AA;

Query Match 100.0%; Score 49; DB 22; Length 25;
Best Local Similarity 23.5%; Pred. NO. 4.8e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

OY 1 XXXXXCXXXXXXCX 17
DB 8 ifccqcnkpsqgicce 24

RESULT 10
AAG64888
ID AAG64888 standard; protein; 25 AA.

Best Local Similarity 70.6%; Pred. No. 4.8e+02;
Matches 12; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCCXXXCCXXXCX 17
Db 8 xfcxcxxxgxcxcx 24

RESULT 13
AAG64891
ID AAG64891 standard; peptide; 25 AA.

AC AAG64891;
DT 15-OCT-2001 (first entry)
DE Human protein related peptide SEQ ID NO: 31.
KW Human; infection; septicemia; drug intoxication; phthisis; cancer;
KW liver dysfunction; immune dysfunction; dysghormonism; antibacterial.
XX Synthetic.

PN JP2001149083-A.
PD 05-JUN-2001.

PF 12-SEP-2000; 2000JP-0276083.
PR 16-SEP-1999; 99JP-0262228.
PA (TAKE) TAKEDA CHEM IND LTD.

DR WPI; 2001-499381/55.
XX New polypeptide and its DNA, useful for diagnosis, treatment and
PT prevention of infection, septicemia, drug intoxication, cancer, liver
PT and immune dysfunction -
XX Disclosure; Page 38; 41pp; Japanese.

XX The present invention provides the protein shown in SEQ ID NO: 3 in the
CC specification. It can be used for the diagnosis, treatment and prevention
CC of opportunistic infections, septicemia, drug intoxication, phthisis,
CC cancers, liver dysfunction, immune dysfunction and dysghormonism. The
CC present sequence is a protein described in the exemplification of the
CC invention.

SQ Sequence 25 AA;

Query Match 100.0%; Score 49; DB 22; Length 25;
Best Local Similarity 64.7%; Pred. No. 4.8e+02;
Matches 11; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCCXXXCCXXXCX 17
Db 8 xfcxcxxxgxcxcx 24

RESULT 14
AAM19888
ID AAM19888 standard; Protein; 25 AA.

AC AAM19888;
DT 12-OCT-2001 (first entry)
DE Peptide #6322 encoded by probe for measuring cervical gene expression.
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.

OS Homo sapiens.
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX

PF 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.

XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 27; SEQ ID No 24714; 487pp; English.

XX The present invention relates to human single exon nucleic acid probes
CC (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 25 AA;

Query Match 100.0%; Score 49; DB 22; Length 25;
Best Local Similarity 23.5%; Pred. No. 4.8e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCCXXXCCXXXCX 17
Db 2 rrpccwqlccwcc 18

RESULT 15
AAM33424
ID AAM33424 standard; Protein; 25 AA.

AC AAM33424;
DT 17-OCT-2001 (first entry)
XX

DE Peptide #7461 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.

OS Homo sapiens.
PN WO200157272-A2.
XX

PD 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0623366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488897/53.

XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 PT
 XX

PS Claim 27; SEQ ID No 33693; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENPs;
 see AAI31315-AAI57546). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders.

XX Sequence 25 AA;
 SQ

Query Match 100.0%; Score 49; DB 22; Length 25;

Best Local Similarity 23.5%; Pred. No. 4.8e+02;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17

Db 2 rrrlccwqicccwawcc 18

RESULT 16

AAU02927

ID AAU02927 standard; Protein; 26 AA.

XX AC AAU02927;

XX DT 11-JUN-1999 (first entry)

XX Fragment of human secreted protein encoded by gene 101.

XX Human; secreted protein; fusin protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

OS Homo sapiens.

XX WO9902546-A1.

XX PD 21-JAN-1999.

XX PF 07-JUL-1998; 98WO-US13684.

XX PR 12-SEP-1997; 97US-0058785.

XX PR 08-JUL-1997; 97US-0051916.

XX PR 08-JUL-1997; 97US-0051918.

XX PR 08-JUL-1997; 97US-0051919.

XX PR 08-JUL-1997; 97US-0051920.

XX PR 08-JUL-1997; 97US-0051925.

XX PR 08-JUL-1997; 97US-0051926.

XX PR 08-JUL-1997; 97US-0051928.

XX PR 08-JUL-1997; 97US-0051929.

PR 08-JUL-1997; 97US-0051930.
 PR 08-JUL-1997; 97US-0051931.
 PR 08-JUL-1997; 97US-0051932.
 PR 08-JUL-1997; 97US-0052732.
 PR 08-JUL-1997; 97US-0052733.
 PR 08-JUL-1997; 97US-0052793.
 PR 08-JUL-1997; 97US-0052795.
 PR 08-JUL-1997; 97US-0052803.
 PR 18-AUG-1997; 97US-0055684.
 PR 18-AUG-1997; 97US-0055722.
 PR 18-AUG-1997; 97US-0055723.
 PR 18-AUG-1997; 97US-0055947.
 PR 18-AUG-1997; 97US-0055948.
 PR 18-AUG-1997; 97US-0055949.
 PR 18-AUG-1997; 97US-0055950.
 PR 18-AUG-1997; 97US-0055953.
 PR 18-AUG-1997; 97US-0055954.
 PR 18-AUG-1997; 97US-0055964.
 PR 18-AUG-1997; 97US-0055984.
 PR 18-AUG-1997; 97US-0056360.
 PR 12-SEP-1997; 97US-0058660.
 PR 12-SEP-1997; 97US-0058661.
 PR 12-SEP-1997; 97US-0058664.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.

XX Brewer LA, Ebner R, Fischer CL, Kyaw H, Lafleur DW, Li Y, Moore PA;
 PI Olsen HS, Rosen CA, Ruben SM, Shi Y, Soppet DR, Zeng Z;
 XX WPI; 1999-120770/10.

XX New isolated human genes and the secreted polypeptides they encode -
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders
 XX
 XX Disclosure; Page 118; 464pp; English.

XX This sequence represents a fragment of a secreted human protein encoded
 CC by the nucleic acid molecule detailed in the descriptor line. The gene
 CC can be used to generate fusion proteins by linking to the gene to a
 CC human immunoglobulin Fc portion (e.g. AAX27302) for increasing the
 CC stability of the fused protein as compared to the human protein only.

CC The invention relates to 123 novel genes and their fragments (nucleic
 CC acid sequences: AAX27311-X27449; amino acid sequences AAY02650-Y02788)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 123
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AAX27311 for described uses).

XX Sequence 26 AA;
 SQ

Query Match 100.0%; Score 49; DB 20; Length 26;

Best Local Similarity 23.5%; Pred. No. 5e+02;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17

Db 5 gtsencrcgcgcgsvcl 21

RESULT 17

AAU05998

ID AAU05998 standard; Peptide; 26 AA.

XX AC AAU05998;

XX DT 24-OCT-2001 (first entry)

XX Cone snail O-superfamily conotoxin, Ts6.2.

XX Cone snail; O-superfamily conotoxin; sodium channel;
 KW demyelinating disease; multiple sclerosis; Huntington's disease;
 KW neuropathy; carpal tunnel syndrome; cardiovascular disorder;
 KW congestive heart failure; cancer; immunosuppression; epilepsy;
 KW asthma; ischaemia; stroke; pain.
 XX Conus tessulatus.
 OS
 XX WO200149312-A2.
 PN
 XX 12-JUL-2001.
 PD
 XX 28-DEC-2000; 2000WO-US35431.
 XX
 XX 30-DEC-1999; 99US-0173754.
 XX
 PR 26-JUN-2000; 2000US-0214263.
 PR
 PR 20-JUL-2000; 2000US-0219440.
 PR
 PR 27-OCT-2000; 2000US-0243412.
 PR
 XX (UTAH) UNIV UTAH RES FOUND.
 DA (COGN-) COGNETIX INC.
 PI Olivera BM, Cartier GE, Watkins M, Hillyard DR, McIntosh JM;
 PI Layer RT, Jones RM;
 DR WPI; 2001-418352/44.
 XX
 XX New O-superfamily polypeptides useful for treating voltage gated ion
 PT channel disorders, including demyelinating diseases i.e. multiple
 PT sclerosis -
 PT
 XX Claim 2; Page 78; 277pp; English.
 PS
 XX The sequence is a cone snail O-superfamily conotoxin peptide.
 CC The peptides are useful for regulating the flow of sodium through
 CC sodium channels in an individual and the treatment or prevention of
 CC disorders associated with voltage gated ion channel disorders,
 CC including demyelinating diseases i.e. multiple sclerosis, optic
 CC neuromyelitis, disseminated encephalomyelitis, adrenoleukodystrophy,
 CC acute transverse myelitis, progressive multifocal leukoencephalopathy,
 CC sub acute sclerosing panencephalomyelitis (SSPE), metachromatic
 CC leukodystrophy, Pelizaeus-Werzbacher disease, spinal cord injury,
 CC botulinum toxin poisoning, Huntington's, compression, entrapment
 CC neuropathies i.e. ulnar nerve palsy, and carpal tunnel syndrome,
 CC cardiovascular disorders, i.e. cardiac arrhythmias and congestive heart
 CC failure, reactive gliosis, hyperglycaemia, immunosuppression, cocaine
 CC addiction, cancer, cognitive dysfunction, neurotransmitter disorders
 CC (i.e. Eaton-Lambert syndrome) and reversal of curare and other
 CC neuromuscular blocking drugs. The neurological disorder is a seizure,
 CC preferably one associated with epilepsy. The neurological disorder is a
 CC neurotoxic injury associated with hypoxia, anoxia or ischaemia. The
 CC neurotoxic injury is associated with stroke, cerebrovascular accident,
 CC brain or spinal cord trauma, myocardial infarct, physical trauma,
 CC drownings, suffocation, perinatal asphyxia or hypoglycaemic events. The
 CC disorder is pain i.e. migraine, acute pain, persistent pain,
 CC neuropathic pain, nociceptive pain. The disorder is inflammation or a
 CC cardiovascular disorder. A conotoxin peptide of is useful to
 CC alleviate pain in a mammal in pain or about to be subjected to a pain
 CC causing event, and to treat disorders associated with radical
 CC depolarisation of excitable membranes by activating a KATP channel, the
 CC disorders include cardiac, ocular and cerebral ischaemia and asthma.
 XX
 XX Sequence 26 AA:
 SQ
 Query Match 100.0%; Score 49; DB 22; Length 26;
 Best Local Similarity 29.4%; Pred. No. 5e+02;
 Matches 5; Conservative 12; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 XXXXCCXXXCCXXXCX 17
 Db ::::||||:|:|:|:
 10 glrgccxgttcfcfclcf 26

RESULT 18

AAU05792
ID AAU05792 standard; Peptide; 27 AA.

XX
AC AAU05792;

XX
XX 24-OCT-2001 (first entry)

XX
DE Cone snail O-superfamily conotoxin, Omaria9.

XX
KW Cone snail; O-superfamily conotoxin; sodium channel;
KW demyelinating disease; multiple sclerosis; Huntington's disease;
KW neuropathy; carpal tunnel syndrome; cardiovascular disorder;
KW congestive heart failure; cancer; immunosuppression; epilepsy;
KW asthma; ischaemia; stroke; pain.

XX
OS Conus omaria.

XX
PN WO200149312-A2.

XX
XX 12-JUL-2001.

XX
PF 28-DEC-2000; 2000WO-US35431.

XX
PR 30-DEC-1999; 99US-0173754.

XX
PR 26-JUN-2000; 2000US-0214263.

XX
PR 20-JUL-2000; 2000US-0219440.

XX
PR 27-OCT-2000; 2000US-0243412.

XX
PA (UTAH) UNIV UTAH RES FOUND.

XX
PA (COGN-) COGNETIX INC.

XX
PI Olivera BM, Cartier GE, Watkins M, Hillyard DR, McIntosh JM;
PI Layer RT, Jones RM;
XX WPI; 2001-418352/44.

XX
PS Claim 2; Page 25; 277pp; English.

CC The sequence is a cone snail O-superfamily conotoxin peptide.
 CC The peptides are useful for regulating the flow of sodium through
 CC sodium channels in an individual and the treatment or prevention of
 CC disorders associated with voltage gated ion channel disorders,
 CC including demyelinating diseases i.e. multiple sclerosis, optic
 CC neuromyelitis, disseminated encephalomyelitis, adrenoleukodystrophy,
 CC acute transverse myelitis, progressive multifocal leukoencephalopathy,
 CC sub acute sclerosing panencephalomyelitis (SSPE), metachromatic
 CC leukodystrophy, Pelizaeus-Werzbacher disease, spinal cord injury,
 CC botulinum toxin poisoning, Huntington's, compression, entrapment
 CC neuropathies i.e. ulnar nerve palsy, and carpal tunnel syndrome,
 CC cardiovascular disorders, i.e. cardiac arrhythmias and congestive heart
 CC failure, reactive gliosis, hyperglycaemia, immunosuppression, cocaine
 CC addiction, cancer, cognitive dysfunction, neurotransmitter disorders
 CC (i.e. Eaton-Lambert syndrome) and reversal of curare and other
 CC neuromuscular blocking drugs. The neurological disorder is a seizure,
 CC preferably one associated with epilepsy. The neurological disorder is a
 CC neurotoxic injury associated with hypoxia, anoxia or ischaemia. The
 CC neurotoxic injury is associated with stroke, cerebrovascular accident,
 CC brain or spinal cord trauma, myocardial infarct, physical trauma,
 CC drownings, suffocation, perinatal asphyxia or hypoglycaemic events. The
 CC disorder is pain i.e. migraine, acute pain, persistent pain,
 CC neuropathic pain, nociceptive pain. The disorder is inflammation or a
 CC cardiovascular disorder. A conotoxin peptide of is useful to
 CC alleviate pain in a mammal in pain or about to be subjected to a pain
 CC causing event, and to treat disorders associated with radical
 CC depolarisation of excitable membranes by activating a KATP channel, the
 CC disorders include cardiac, ocular and cerebral ischaemia and asthma.

XX
SQ Sequence 27 AA;

Query Match 100.0%; Score 49; DB 22; Length 27;
Best Local Similarity 29.4%; Pred. No. 5.1e+02;
Matches 5; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCXXXXXXCX 17
J:::||||::|::|:
Db 11 xifqncghlfcvvcv 27

RESULT 19
ABG13085
ID ABG13085 standard; Protein; 28 AA.
XX

ABG13085;

13-FEB-2002 (first entry)

Novel human diagnostic protein #13076.

Human; chromosome mapping; gene mapping; gene therapy; forensic;
food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS77272.

PT New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits and to assess
biodiversity -

PS Claim 20; SEQ ID No 43444; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 28 AA;

Query Match 100.0%; Score 49; DB 22; Length 28;
Best Local Similarity 23.5%; Pred. No. 5.2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCXXXXXXCX 17
J:::||||::|::|:
Db 3 vnrssccccccccccc 19

RESULT 20
AAR53575
ID AAR53575 standard; peptide; 29 AA.
XX

AC AAR53575;

XX 29-NOV-1994 (first entry)

DE Spider venom calcium channel blocking peptide KJ-1.

KW Spider venom; calcium channel blocking protein; calcium-antagonist;
KW angina; hypertension; cardiomyopathy; pesticide.

OS Heteropoda venatoria.

PN WO9410195-A.

PD 11-MAY-1994.

XX 16-AUG-1993; 93WO-US07555.

XX 30-OCT-1992; 92US-0970333.

XX (PFIZ) PFIZER INC.

XX Kelbaugh PR, Saccomano NA, Volkman RA;

XX WPI; 1994-167384/20.

PT Calcium channel-blocking polypeptide(s) from heteropoda venatoria
PT venom - used to treat e.g. angina, hypertension,
PT cardiomyopathies, etc. and for invertebrate pest control

XX Claim 21; Page 22; 31pp; English.

XX The peptide is useful in blocking Ca channels in cells, such as
CC cells in the nervous system of a mammal, in the treatment of Ca
CC channel-mediated diseases and conditions (e.g. angina, hypertension,
CC cardiomyopathy, supraventricular arrhythmias, esophageal achalasia,
CC premature labor and Raynaud's disease. The peptides are obtained
CC from the spider through the process of milking by electrical
CC stimulation.

XX Sequence 29 AA;

Query Match 100.0%; Score 49; DB 15; Length 29;
Best Local Similarity 23.5%; Pred. No. 5.4e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCXXXXXXCX 17
J:::||||::|::|:
Db 11 dtskdccgyvhlwck 27

RESULT 21
AAW51455
ID AAW51455 standard; peptide; 29 AA.
XX

AC AAW51455;

XX

XX DNA encoding conotoxin MrVIB precursor polypeptides which are useful as
PT tools for immobilizing skeletal muscle, as reagents for investigative
PT voltage-dependent calcium ion channels and for drug screening -

CC sub acute sclerosing panencephalomyelitis (SSPE), metachromatic
 CC leukodystrophy, Pelizaeus-Werzbacher disease, spinal cord injury,
 CC botulinum toxin poisoning, Huntington's, compression, entrapment
 CC neuropathies i.e. ulnar nerve palsy, and carpal tunnel syndrome,
 CC cardiovascular disorders, i.e. cardiac arrhythmias and congestive heart
 CC failure, reactive gliosis, hyperglycaemia, immunosuppression, cocaine
 CC addiction, cancer, cognitive dysfunction, neurotransmitter disorders
 CC (i.e. Eaton-Lambert syndrome) and reversal of curare and other
 CC neuromuscular blocking drugs. The neurological disorder is a seizure,
 CC preferably one associated with epilepsy. The neurological disorder is a
 CC neurotoxic injury associated with hypoxia, anoxia or ischaemia. The
 CC neurotoxic injury is associated with stroke, cerebrovascular accident,
 CC brain or spinal cord trauma, myocardial infarct, physical trauma,
 CC drownings, suffocation, perinatal asphyxia or hypoglycaemic events. The
 CC disorder is pain i.e. migraine, acute pain, persistent pain,
 CC neuropathic pain, nociceptive pain. The disorder is inflammation or a
 CC cardiovascular disorder. A conotoxin peptide of is useful to
 CC alleviate pain in a mammal in pain or about to be subjected to a pain
 CC causing event, and to treat disorders associated with radical
 CC depolarisation of excitable membranes by activating a KATP channel, the
 CC disorders include cardiac, ocular and cerebral ischaemia and asthma.

Sequence 29 AA;

Query Match 100.0%; Score 49; DB 22; Length 29;
 Best Local Similarity 35.3%; Pred. No. 5.4e+02;
 Matches 6; Conservative 11; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCCX 17
 I:::||||:|:::|
 Db 13 xifqncrcgxcnvcfvcv 29

RESULT 26

AAU05789
 ID AAU05789 standard; Peptide; 29 AA.

AC AAU05789;

DT 24-OCT-2001 (first entry)

DE Cone snail O-superfamily conotoxin, Delta-GmVIA [F15V].

KW Cone snail; O-superfamily conotoxin; sodium channel;
 KW demyelinating disease; multiple sclerosis; Huntington's disease;
 KW neuropathy; carpal tunnel syndrome; cardiovascular disorder;
 KW congestive heart failure; cancer; immunosuppression; epilepsy;
 KW asthma; ischaemia; stroke; pain.

OS Conus gloriamaris.

PN WO200149312-A2.

PD 12-JUL-2001.

PF 28-DEC-2000; 2000WO-US35431.

PR 30-DEC-1999; 99US-0173754.

PR 26-JUN-2000; 2000US-0214263.

PR 20-JUL-2000; 2000US-0219440.

PR 27-OCT-2000; 2000US-0243412.

PA (UTAH) UNIV UTAH RES FOUND.

PA (COGN-) COGNETIX INC.

PI Olivera BM, Cartier GE, Watkins M, Hillyard DR, McIntosh JM;

PI Layer RT, Jones RM;

XX WPI; 2001-418352/44.

XX New O-superfamily polypeptides useful for treating voltage gated ion
 channel disorders, including demyelinating diseases i.e. multiple

PT sclerosis -

PS Claim 2; Page 24; 277pp; English.

XX The sequence is a cone snail O-superfamily conotoxin peptide.
 CC The peptides are useful for regulating the flow of sodium through
 CC sodium channels in an individual and the treatment or prevention of
 CC disorders associated with voltage gated ion channel disorders,
 CC including demyelinating diseases i.e. multiple sclerosis, optic
 CC neuromyelitis, disseminated encephalomyelitis, adrenoleukodystrophy,
 CC acute transverse myelitis, progressive multifocal leukoencephalopathy,
 CC sub acute sclerosing panencephalomyelitis (SSPE), metachromatic
 CC leukodystrophy, Pelizaeus-Werzbacher disease, spinal cord injury,
 CC botulinum toxin poisoning, Huntington's, compression, entrapment
 CC neuropathies i.e. ulnar nerve palsy, and carpal tunnel syndrome,
 CC cardiovascular disorders, i.e. cardiac arrhythmias and congestive heart
 CC failure, reactive gliosis, hyperglycaemia, immunosuppression, cocaine
 CC addiction, cancer, cognitive dysfunction, neurotransmitter disorders
 CC (i.e. Eaton-Lambert syndrome) and reversal of curare and other
 CC neuromuscular blocking drugs. The neurological disorder is a seizure,
 CC preferably one associated with epilepsy. The neurological disorder is a
 CC neurotoxic injury associated with hypoxia, anoxia or ischaemia. The
 CC neurotoxic injury is associated with stroke, cerebrovascular accident,
 CC brain or spinal cord trauma, myocardial infarct, physical trauma,
 CC drownings, suffocation, perinatal asphyxia or hypoglycaemic events. The
 CC disorder is pain i.e. migraine, acute pain, persistent pain,
 CC neuropathic pain, nociceptive pain. The disorder is inflammation or a
 CC cardiovascular disorder. A conotoxin peptide of is useful to
 CC alleviate pain in a mammal in pain or about to be subjected to a pain
 CC causing event, and to treat disorders associated with radical
 CC depolarisation of excitable membranes by activating a KATP channel, the
 CC disorders include cardiac, ocular and cerebral ischaemia and asthma.

XX Sequence 29 AA;

Query Match 100.0%; Score 49; DB 22; Length 29;

Best Local Similarity 41.2%; Pred. No. 5.4e+02;

Matches 7; Conservative 10; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCCX 17

I:::||||:|:::|

Db 13 xifqncrcgxcnvcfvcv 29

RESULT 27

AAU05790

ID AAU05790 standard; Peptide; 29 AA.

XX AAU05790;

DT 24-OCT-2001 (first entry)

DE Cone snail O-superfamily conotoxin, Delta-GmVIA [F27Y].

KW Cone snail; O-superfamily conotoxin; sodium channel;
 KW demyelinating disease; multiple sclerosis; Huntington's disease;
 KW neuropathy; carpal tunnel syndrome; cardiovascular disorder;
 KW congestive heart failure; cancer; immunosuppression; epilepsy;
 KW asthma; ischaemia; stroke; pain.

OS Conus gloriamaris.

PN WO200149312-A2.

PD 12-JUL-2001.

PF 28-DEC-2000; 2000WO-US35431.

PR 30-DEC-1999; 99US-0173754.

PR 26-JUN-2000; 2000US-0214263.

PR 20-JUL-2000; 2000US-0219440.

PR 27-OCT-2000; 2000US-0243412.

```

PD 11-MAY-1994.
XX
XX 16-AUG-1993; 93WO-US07555.
XX
XX 30-OCT-1992; 92US-0970333.
XX
XX (PFIZ ) PFIZER INC.
XX
XX Kelbaugh PR, Saccomano NA, Volkmann RA;
XX WPI; 1994-167384/20.
XX
XX Calcium channel-blocking polypeptide(s) from heteropoda venatoria
PT venom - used to treat e.g. angina, hypertension,
PT cardiomyopathies, etc. and for invertebrate pest control
XX
XX Claim 1; Page 19; 31pp; English.
XX
XX The peptide is useful in blocking Ca channels in cells, such as
XX cells in the nervous system of a mammal, in the treatment of Ca
CC channel-mediated diseases and conditions (e.g. angina, hypertension)
CC cardiomyopathy, supraventricular arrhythmias, esophageal achalasia,
CC premature labor and Raynaud's disease. The peptides are obtained
CC from the spider through the process of milking by electrical
CC stimulation.
XX
XX Sequence 30 AA;
XX
XX Query Match 100.0%; Score 49; DB 15; Length 30;
XX Best Local Similarity 23.5%; Pred. No. 5.5e+02;
XX Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps
XX
XX QY 1 XXXXXCCXXXCCXXCX 17
XX ::::|:::|:::|:
XX DB 11 dsnadcegyvrclwck 27
XX
XX RESULT 29
XX AAR53577
XX ID AAR53577 standard; peptide; 30 AA.
XX
XX AAR53577;
XX
XX 29-NOV-1994 (first entry)
XX
XX Spider venom calcium channel blocking peptide KJ-6.
XX
XX Spider venom; calcium channel blocking protein; calcium-antagonist;
XX angina; hypertension; cardiomyopathy; pesticide.
XX
XX Heteropoda venatoria.
XX
XX WO9410195-A.
XX
XX 11-MAY-1994.
XX
XX 16-AUG-1993; 93WO-US07555.
XX
XX 30-OCT-1992; 92US-0970333.
XX
XX (PFIZ ) PFIZER INC.
XX
XX Kelbaugh PR, Saccomano NA, Volkmann RA;
XX WPI; 1994-167384/20.
XX
XX Calcium channel-blocking polypeptide(s) from heteropoda venatoria
PT venom - used to treat e.g. angina, hypertension,
PT cardiomyopathies, etc. and for invertebrate pest control
XX
XX Claim 29; Page 23; 31pp; English.
XX

```

CC The peptide is useful in blocking Ca channels in cells, such as
CC cells in the nervous system of a mammal, in the treatment of Ca
CC channel-mediated diseases and conditions (e.g. angina, hypertension,
CC cardiomyopathy, supraventricular arrhythmias, esophageal achalasia,
CC premature labor and Raynaud's disease. The peptides are obtained
CC from the spider through the process of milking by electrical
CC stimulation.
XX
SQ Sequence 30 AA;

Query Match 100.0%; Score 49; DB 15; Length 30;
Best Local Similarity 23.5%; Pred. No. 5.5e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCXXXCXXXCX 17
Db 11 dtnadccgyvcrllwck 27

RESULT 30
AAR63751
D AAR63751 standard; Peptide; 30 AA.

XX
AC AAR63751;
XX
DT 19-MAY-1995 (first entry)
XX
DE Outward K channel inhibitor from the spider H. venatoria, peptide 1.
XX
KW Venom; spider; Heteropoda venatoria; Olios fasciculatus; inhibitor;
KW potassium channel; chromatography; disulphide bridge; S-pyridilation;
KW activating agent; cardiac cell; neuron; cardiac; arrhythmia; treatment;
KW learning; memory; disorder; Alzheimer's disease; Parkinson's disease;
KW multiple sclerosis; schizophrenia; epilepsy; stroke; muscle spasticity;
KW insecticidal agent; insecticide.
XX

OS Heteropoda venatoria.

XX WO9421278-A.

XX 29-SEP-1994.

XX 14-MAR-1994; 94WO-US02750.

XX 18-MAR-1993; 93US-0033388.

XX (NPSP-) NPS PHARM INC.

XX Mueller AL, Sanguinetti MC;

XX WPI; 1994-316654/39.

XX New transient outward potassium channel inhibitors from spider
PT venom - is used to treat cardiac arrhythmias and disorders of
PT learning and memory such as Alzheimer's disease
XX

PS Example 2; Page 18; 57pp; English.

XX The amino acid sequence of a peptide isolated from the venom of the
CC spider Heteropoda venatoria which has an inhibitory effect on transient
CC outward potassium channels (TOPCs). The compound was isolated from the
CC crude venom by conventional chromatographic methods e.g. reverse-phase
CC HPLC. The sequence was determined using a pulse-liquid sequenator. The
CC protein has a molecular mass of 3599 as calculated by ion-spray mass
CC spectroscopy and a calculated pI of 5.41. The protein is thought to
CC contain 3 disulphide bridges which can be broken by S-pyridilation prior
CC to sequencing. The TOPC inhibitor (TOPCI) can be used to screen for
CC TOPC activating agents (TOPCAs). The TOPCIs selectively block the
CC outward currents in cardiac cells and in neurons. They can be used
CC in the treatment of cardiac arrhythmias and in the treatment of
CC disorders of learning and memory such as Alzheimer's disease.
CC Parkinson's disease, multiple sclerosis, schizophrenia, epilepsy, stroke

CC and muscle spasticity. The compounds can also be used as insecticidal
CC agents. The compound is related in activity to AAR63751-3.
XX
SQ Sequence 30 AA;

Query Match 100.0%; Score 49; DB 15; Length 30;
Best Local Similarity 23.5%; Pred. No. 5.5e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCXXXCXXXCX 17
Db 11 dtnadccgyvcrllwck 27

RESULT 31
AAO10819
ID AAO10819 standard; Protein; 30 AA.

XX
AC AAO10819;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 24711.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX

OS Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US04927.

XX 28-FEB-2000; 2000US-0515126.

XX 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

XX N-PSDB; AAI90750.

XX Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -

XX Claim 20; SEQ ID NO 24711; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.

XX Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 30 AA;

Query Match 100.0%; Score 49; DB 22; Length 30;

Best Local Similarity 29.4%; Pred. No. 5.5e+02;
Matches 5; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCCXXXCCXXXCX 17
Db 1 cccccccccccccccc 17

RESULT 32

AAU05791
ID AAU05791 standard; Peptide; 30 AA.

XX
AC AAU05791;

XX
DT 24-OCT-2001 (first entry)

XX
DE Cone snail O-superfamily conotoxin propeptide, Omaria9.

XX
KW Cone snail; O-superfamily conotoxin; sodium channel;
demylinating disease; multiple sclerosis; Huntington's disease;
neuropathy; carpal tunnel syndrome; cardiovascular disorder;
congestive heart failure; cancer; immunosuppression; epilepsy;
asthma; ischaemia; stroke; pain.

XX
OS Conus omaria.

XX
PN WO200149312-A2.

XX
PD 12-JUL-2001.

XX
PF 28-DEC-2000; 2000WO-US35431.

XX
PR 30-DEC-1999; 99US-0173754.

XX
PR 26-JUN-2000; 2000US-0214263.

XX
PR 20-JUL-2000; 2000US-0219440.

XX
PR 27-OCT-2000; 2000US-0243412.

XX
PA (UTAH) UNIV UTAH RES FOUND.

XX
PA (COGN-) COGNETIX INC.

XX
PI Olivera BM, Cartier GE, Watkins M, Hillyard DR, McIntosh JM;

XX
PI Layer RT, Jones RM;

XX
DR WPI; 2001-418352/44.

XX
DR N-PSDB; AAS10887.

XX
XX New O-superfamily polypeptides useful for treating voltage gated ion

XX
XX channel disorders, including demyelinating diseases i.e. multiple

XX
XX sclerosis -

XX
XX Claim 15; Page 25; 277pp; English.

XX
XX The sequence is a cone snail O-superfamily conotoxin propeptide.

XX
XX The peptides are useful for regulating the flow of sodium through

XX
XX sodium channels in an individual and the treatment or prevention of

XX
XX disorders associated with voltage gated ion channel disorders,

XX
XX including demyelinating diseases i.e. multiple sclerosis, optic

XX
XX neuroneitis, disseminated encephalomyelitis, adrenoleukodystrophy,

XX
XX acute transverse myelitis, progressive multifocal leukoencephalopathy,

XX
XX sub acute sclerosing panencephalomyelitis (SSP), metachromatic

XX
XX leukodystrophy, Pelizaeus-Merzbacher disease, spinal cord injury,

XX
XX botulinum toxin poisoning, Huntington's, compression, entrapment

CC disorder is pain i.e. migraine, acute pain, persistent pain,
CC neuropathic pain, nociceptive pain. The disorder is inflammation or a
CC cardiovascular disorder. A conotoxin peptide of is useful to
CC alleviate pain in a mammal in pain or about to be subjected to a pain
CC causing event, and to treat disorders associated with radical
CC depolarisation of excitable membranes by activating a KATP channel, the
CC disorders include cardiac, ocular and cerebral ischaemia and asthma.

XX
SQ Sequence 30 AA;

Query Match 100.0%; Score 49; DB 22; Length 30;
Best Local Similarity 23.5%; Pred. No. 5.5e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCCXXXCCXXXCX 17
Db 14 pifqncchglfcvllcv 30

RESULT 33

AAU53574
ID AAU53574 standard; peptide; 31 AA.

XX
AC AAR53574;

XX
DT 29-NOV-1994 (first entry)

XX
DE Spider venom calcium channel blocking peptide AU-5C.

XX
KW Spider venom; calcium channel blocking protein; calcium-antagonist;
angina; hypertension; cardiomyopathy; pesticide.

XX
OS Heteropoda venatoria.

XX
PN WO9410195-A.

XX
PD 11-MAY-1994.

XX
PF 16-AUG-1993; 93WO-US07555.

XX
PR 30-OCT-1992; 92US-0970333.

XX
PA (PFIZ) PFIZER INC.

XX
PI Kelbaugh PR, Saccomano NA, Volkmann RA;

XX
XX WPI; 1994-167384/20.

XX
XX Calcium channel-blocking polypeptide(s) from heteropoda venatoria

XX
XX venom - used to treat e.g. angina, hypertension,

XX
XX cardiomyopathies, etc. and for invertebrate pest control

XX
PS Claim 17; Page 21; 3lpp; English.

XX
XX The peptide is useful in blocking Ca channels in cells, such as
XX cells in the nervous system of a mammal, in the treatment of Ca
XX channel-mediated diseases and conditions (e.g. angina, hypertension,
XX cardiomyopathy, supraventricular arrhythmias, esophageal achalasia,
XX premature labor and Raynaud's disease. The peptides are obtained
XX from the spider through the process of milking by electrical
XX stimulation.

XX
SQ Sequence 31 AA;

Query Match 100.0%; Score 49; DB 15; Length 31;
Best Local Similarity 23.5%; Pred. No. 5.7e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCCXXXCCXXXCX 17
Db 10 sthadcccgfcklwr 26

RESULT 34
AAR53578
ID AAR53578 standard; peptide; 31 AA.
XX AC AAR53578;
XX DT 29-NOV-1994 (first entry)
XX DE Spider venom calcium channel blocking peptide KJ-7.
XX KW Spider venom; calcium channel blocking protein; calcium-antagonist;
XX KW angina; hypertension; cardiomyopathy; pesticide.
OS Heteropoda venatoria.
XX PN W09410195-A.
XX PD 11-MAY-1994.
XX 16-AUG-1993; 93WO-US07555.
XX 30-OCT-1992; 92US-0970333.
XX (PFIZ) PFIZER INC.
XX Kelbaugh PR, Saccomano NA, Volkmann RA;
XX WPI; 1994-167384/20.
XX Calcium channel-blocking polypeptide(s) from heteropoda venatoria
PT venom - used to treat e.g. angina, hypertension,
PT cardiomyopathies, etc. and for invertebrate pest control
XX
XX Claim 33; Page 23; 31pp; English.
XX The peptide is useful in blocking Ca channels in cells, such as
XX cells in the nervous system of a mammal, in the treatment of Ca
XX channel-mediated diseases and conditions (e.g. angina, hypertension,
XX cardiomyopathy, supraventricular arrhythmias, esophageal achalasia,
XX premature labor and Raynaud's disease. The peptides are obtained
XX from the spider through the process of milking by electrical
XX stimulation.
XX
XX Sequence 31 AA;
Query Match 100.0%; Score 49; DB 15; Length 31;
Best Local Similarity 23.5%; Pred. No. 5.7e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;
QY 1 XXXXXCCXXXCCXXXCX 17
Db 10 sthaddcegfcklwr 26
RESULT 35
AAR63752
ID AAR63752 standard; Peptide; 31 AA.
XX AC AAR63752;
XX DT 19-MAY-1995 (first entry)
XX DE Outward K channel inhibitor from the spider H. venatoria, peptide 2.
XX KW Venom; spider; Heteropoda venatoria; Olios fasciculatus; inhibitor;
XX KW potassium channel; chromatography; disulphide bridge; S-Pyridilation;
XX KW activating agent; cardiac cell; neuron; cardiac; arrhythmia; treatment;
XX KW learning; memory; disorder; Alzheimer's disease; Parkinson's disease;
XX KW multiple sclerosis; schizophrenia; epilepsy; stroke; muscle spasticity;
XX KW insecticidal agent; insecticide.

XX OS Heteropoda venatoria.
XX PN W09421278-A.
XX PD 29-SEP-1994.
XX 14-MAR-1994; 94WO-US02750.
XX 18-MAR-1993; 93US-0033388.
XX (NPS-) NPS PHARM INC.
XX Mueller AL, Sanguinetti MC;
XX WPI; 1994-316654/39.
XX New transient outward potassium channel inhibitors from spider
PT venom - is used to treat cardiac arrhythmias and disorders of
PT learning and memory such as Alzheimer's disease
XX Example 4; Page 20; 57pp; English.
XX The amino acid sequence of a peptide isolated from the venom of the
CC spider Heteropoda venatoria which has an inhibitory effect on transient
CC outward potassium channels (TOPCs). The compound was isolated from the
CC crude venom by conventional chromatographic methods e.g. reverse-phase
CC HPLC. The sequence was determined using a pulse-liquid sequenator. The
CC protein has a molecular mass of 3599 as calculated by ion-spray mass
CC spectroscopy and a calculated pI of 5.41. The protein is thought to
CC contain 3 disulphide bridges which can be broken by S-pyridilation prior
CC to sequencing. The TOPC inhibitor (TOPCI) can be used to screen for
CC TOPC activating agents (TOPCAs). The TOPCIs selectively block the
CC outward currents in cardiac cells and in neurons. They can be used
CC in the treatment of cardiac arrhythmias and in the treatment of
CC Parkinson's disease, multiple sclerosis, schizophrenia, epilepsy, stroke
CC and muscle spasticity. The compounds can also be used as insecticidal
XX agents. The compound is related in activity to AAR63751-3.
XX Sequence 31 AA;

Query Match 100.0%; Score 49; DB 15; Length 31;
Best Local Similarity 23.5%; Pred. No. 5.7e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
Db 10 sthaddcegfcklwr 26

RESULT 36
AAW51456
ID AAW51456 standard; peptide; 31 AA.
XX AC AAW51456;

XX DT 02-SEP-1998 (first entry)
XX DE Analgesic peptide from venom of Grammostola spatulata.
XX KW Analgesic; Chilean pink tarantula spider; pain; venom; cancer;
XX cross-tolerance.
XX OS Grammostola spatulata.

XX PN US5776896-A.
XX PD 07-JUL-1998.
XX 30-DEC-1996; 96US-0775476.

PS Claim 1; Column 11-12; 7pp; English.

XX The present sequence represents a purified peptide from Grammostola

CC spatulata spider venom. The peptide is referred to as alpha-mechanotoxin

CC or GSAF II. The peptide is useful for mediating hypotonic cell swelling

CC induced calcium increase in cells, and is especially useful for treating

CC cardiac arrhythmia.

XX Sequence 31 AA;

SQ

Query Match 100.0%; Score 49; DB 19; Length 31;
Best Local Similarity 23.5%; Pred. NO. 5.7e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps

QY 1 XXXXXCXXXXXXXCX 17
 |:::||||::|::|:

Db 10 deerkceglvrclwck 26

RESULT 38

AAY45128

ID AAY45128 standard; peptide; 31 AA.

XX AC

XX AAY45128;

XX DT

XX 22-DEC-1999 (first entry)

DE Chilean pink tarantula spider venom mechanotoxin peptide.

XX KW

KW Chilean pink tarantula spider; Grammostola spatulata; venom;

KW stretch-activated channel block; cell membrane; mechanotoxin;

KW cardiac ventricular rhythm disturbance; GSAF II; antiarrhythmic;

KW hypotonic cell swelling.

XX OS

OS Grammostola spatulata.

XX US5968838-A.

PN 19-OCT-1999.

PD

PF 21-JAN-1998; 98US-0010204.

PP

PR 03-JAN-1996; 96US-0009580.

PR 30-DEC-1996; 96US-0775477.

XX (ZENE) ZENECA LTD.

XX PA

XX Sachs F, Lampe RA;

PI PI

PI WPI: 1999-600830/51.

DR XX

XX Identifying a compound which mimics a peptide on a stretch-activated

PT channel -

PT

PS Claim 1; Column 11-12; 7pp; English.

PS

XX A method has been developed of identifying a compound which mimics the

CC binding activity of a peptide with the present 31 amino acid sequence on

CC a stretch-activated channel. The method comprises: (a) incorporating a

CC detectable label into the peptide; (b) binding the labelled peptide to a

CC stretch-activated channel in a sample; (c) adding a test compound to the

CC sample; and (d) detecting the binding activity of the test compound.

CC The peptide is isolated from the venom of the Chilean pink tarantula

CC spider (Grammostola spatulata), and is designated mechanotoxin or

CC GSAF II. The peptide and its related compounds can be used for the

CC treatment of arrhythmia.

XX Sequence 31 AA;

SQ

Query Match 100.0%; Score 49; DB 20; Length 31;
Best Local Similarity 23.5%; Pred. NO. 5.7e+02;

XX Cone snail; alpha-conotoxin; venom; disulphide bond; mood disorder;
KW

PD 03-2

XX 28-JAN-2000; 2000WO-US01979.
PF N-PSDB; AAA89450.
XX
PR 29-JAN-1999; 99US-0118381.
XX
XX (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
PI Watkins M, Olivera BM, Hillyard DR, McIntosh JM, Jones RM;
XX
XX WPI: 2000-505965/45.
DR N-PSDB; AAA89449.
XX
XX alpha-conotoxin polypeptides derived from the venom of cone snails
PT useful e.g. as neuromuscular blocking agents for use in surgery and for
PT treating unipolar depression -
XX
XX Claim 39; Page 46; 229pp; English.
XX
XX The present invention relates to a number of alpha-conotoxin peptides and
CC their coding sequences from a number of different species of cone snail.
CC These peptides are found in minute quantities in the venom of the snails,
CC and are targeted at the neuronal nicotinic acetylcholine receptors of the
CC nervous system. They usually contain two disulphide bonds, which give
CC them defined conformations, a rarity in molecules this small. The
CC alpha-conotoxins can be used as neuromuscular blocking agents in surgery,
CC and for treating disorders regulated at the neuronal nicotinic
CC acetylcholine receptors, including cardiovascular disorders, gastric
CC motility disorders, urinary incontinence, nicotine addiction, mood
CC disorders such as bipolar disorder, unipolar depression, dysthymia and
CC seasonal affective disorder, and small cell lung carcinoma.
XX
XX Sequence 35 AA;
SQ

Query Match 100.0%; Score 49; DB 21; Length 35;
Best Local Similarity 23.5%; Pred. No. 6.2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;
QY 1 XXXXXCCXXXCCXXCX 17
DB 19 firdecdsrccgknc 35

RESULT 44
AAB21554
ID AAB21554 standard; Peptide; 35 AA.
XX
XX AAB21554;
XX
XX 19-JAN-2001 (first entry)
XX
XX Cone snail alpha-conotoxin SEQ ID NO: 236.
DE
XX
XX Cone snail; alpha-conotoxin; venom; disulphide bond; mood disorder;
KW neuronal nicotinic acetylcholine receptor; cardiovascular disorder;
KW gastric motility disorder; urinary incontinence; nicotine addiction;
KW small cell lung carcinoma.
XX
XX Conus imperialis.
OS
XX
XX WO2000044776-A1.
PN
XX
XX 03-AUG-2000.
PD
XX
XX 28-JAN-2000; 2000WO-US01979.
PF
XX
XX 29-JAN-1999; 99US-0118381.
PR
XX
XX (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
XX Watkins M, Olivera BM, Hillyard DR, McIntosh JM, Jones RM;

XX WPI: 2000-505965/45.
DR N-PSDB; AAA89450.
XX
XX alpha-conotoxin polypeptides derived from the venom of cone snails
PT useful e.g. as neuromuscular blocking agents for use in surgery and for
PT treating unipolar depression -
XX
XX Claim 39; Page 46; 229pp; English.
XX
XX The present invention relates to a number of alpha-conotoxin peptides and
CC their coding sequences from a number of different species of cone snail.
CC These peptides are found in minute quantities in the venom of the snails,
CC and are targeted at the neuronal nicotinic acetylcholine receptors of the
CC nervous system. They usually contain two disulphide bonds, which give
CC them defined conformations, a rarity in molecules this small. The
CC alpha-conotoxins can be used as neuromuscular blocking agents in surgery,
CC and for treating disorders regulated at the neuronal nicotinic
CC acetylcholine receptors, including cardiovascular disorders, gastric
CC motility disorders, urinary incontinence, nicotine addiction, mood
CC disorders such as bipolar disorder, unipolar depression, dysthymia and
CC seasonal affective disorder, and small cell lung carcinoma.
XX
XX Sequence 35 AA;
SQ

Query Match 100.0%; Score 49; DB 21; Length 35;
Best Local Similarity 23.5%; Pred. No. 6.2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;
QY 1 XXXXXCCXXXCCXXCX 17
DB 19 ivrraccsdrccrcg 35

RESULT 45
AAB21555
ID AAB21555 standard; Peptide; 35 AA.
XX
XX AAB21555;
XX
XX 19-JAN-2001 (first entry)
XX
XX Cone snail alpha-conotoxin SEQ ID NO: 238.
DE
XX
XX Cone snail; alpha-conotoxin; venom; disulphide bond; mood disorder;
KW neuronal nicotinic acetylcholine receptor; cardiovascular disorder;
KW gastric motility disorder; urinary incontinence; nicotine addiction;
KW small cell lung carcinoma.
XX
XX Conus regius.
OS
XX
XX WO2000044776-A1.
PN
XX
XX 03-AUG-2000.
PD
XX
XX 28-JAN-2000; 2000WO-US01979.
PF
XX
XX 29-JAN-1999; 99US-0118381.
PR
XX
XX (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
XX Watkins M, Olivera BM, Hillyard DR, McIntosh JM, Jones RM;
PI
XX
XX WPI: 2000-505965/45.
DR N-PSDB; AAA89451.
XX
XX alpha-conotoxin polypeptides derived from the venom of cone snails
PT useful e.g. as neuromuscular blocking agents for use in surgery and for
PT treating unipolar depression -
XX
XX Claim 39; Page 46; 229pp; English.

```

Query Match      100.0%; Score 49; DB 21; Length 35;
Best Local Similarity 23.5%; Pred. No. 6.2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXGCCXXXXXXCX 17
    ::::|:::|:::|:::|
Db 19 flrdvccrthactlca 35

```

RESULT 47
AAB211557
ID AAB21557 standard; Peptide; 35 AA.
XX
AC
XX AAB21557;
XX
DT 19-JAN-2001 (first entry)

| | | | |
|----|-------------|-------------------------|---|
| DE | Cone snail | alpha-conotoxin | SEQ ID NO: 242. |
| XX | | | |
| KW | Cone snail; | alpha-conotoxin; | venom; disulphide bond; mood disorder; |
| KW | neuronal | nicotinic acetylcholine | receptor; cardiovascular disorder; |
| KW | gastric | motility disorder; | urinary incontinence; nicotine addiction; |
| KW | small cell | lung carcinoma. | |

29-JAN-1999; 9905-0118361.
XX
XX
PA (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
XX
PI Watkins M, Olivera BM, Hillvard DR, McIntosh JM, Jones RM:

XX
DR WPI; 2000-505965/45.
DR N-PSDB; AAA89453.
XX
XX
PT alpha-conotoxin polypeptides derived from the venom of cone snails
PT useful e.g. as neuromuscular blocking agents for use in surgery and for
PT treating unipolar depression -
XX
PS Claim 39; Page 47; 229pp; English.

CC acetylcholine receptors, including cardiovascular disorders, gastric
CC motility disorders, urinary incontinence, nicotine addiction, mood
CC disorders such as bipolar disorder, unipolar depression, dysthymia and
CC seasonal affective disorder, and small cell lung carcinoma.
XX
XX
SQ Sequence 35 AA;
Query Match 100.0%; Score 49; DB 21; Length 35;

Query Match 100.0%; Score 49; DB 21; Length 35;

Best Local Similarity 23.5%; Pred. No. 6.2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
Db 19 firdycrrppctliog 35

RESULT 48

AAB21559
ID AAB21559 standard; Peptide; 35 AA.
XX AC AAB21559;
XX AC AAB21559;
XX AC AAB21559;
DT 19-JAN-2001 (first entry)
DE Cone snail alpha-conotoxin SEQ ID NO: 246.
XX Cone snail; alpha-conotoxin; venom; disulphide bond; mood disorder;
KW neuronal nicotinic acetylcholine receptor; cardiovascular disorder;
KW gastric motility disorder; urinary incontinence; nicotine addiction;
KW small cell lung carcinoma.
XX Conus regius.
OS WO200044776-A1.
XX 03-AUG-2000.
XX 28-JAN-2000; 2000WO-US01979.
XX 29-JAN-1999; 99US-0118381.
XX (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX Watkins M, Olivera BM, Hillyard DR, McIntosh JM, Jones RM;
PI WPI: 2000-505965/45.
DR N-PSDB; AAA89455.
XX alpha-conotoxin polypeptides derived from the venom of cone snails
PT useful e.g. as neuromuscular blocking agents for use in surgery and for
PT treating unipolar depression -
XX Claim 39; Page 47; 229pp; English.

XX The present invention relates to a number of alpha-conotoxin peptides and
CC their coding sequences from a number of different species of cone snail.
CC These peptides are found in minute quantities in the venom of the snails,
CC and are targeted at the neuronal nicotinic acetylcholine receptors of the
CC nervous system. They usually contain two disulphide bonds, which give
CC them defined conformations, a rarity in molecules this small. The
CC alpha-conotoxins can be used as neuromuscular blocking agents in surgery,
CC acetylcholine receptors, including cardiovascular disorders, gastric
CC motility disorders, urinary incontinence, nicotine addiction, mood
CC disorders such as bipolar disorder, unipolar depression, dysthymia and
CC seasonal affective disorder, and small cell lung carcinoma.
XX Sequence 35 AA;

Query Match 100.0%; Score 49; DB 21; Length 35;
Best Local Similarity 23.5%; Pred. No. 6.2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
Db 19 vvirgcccprcawrog 35

RESULT 49

AAB21560
ID AAB21560 standard; Peptide; 35 AA.
XX AC AAB21560;
XX AC AAB21560;
DT 19-JAN-2001 (first entry)
DE Cone snail alpha-conotoxin SEQ ID NO: 248.
XX Cone snail; alpha-conotoxin; venom; disulphide bond; mood disorder;
KW neuronal nicotinic acetylcholine receptor; cardiovascular disorder;
KW gastric motility disorder; urinary incontinence; nicotine addiction;
KW small cell lung carcinoma.
XX Conus regius.
OS WO200044776-A1.
XX 03-AUG-2000.
XX 28-JAN-2000; 2000WO-US01979.
XX 29-JAN-1999; 99US-0118381.
XX (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX Watkins M, Olivera BM, Hillyard DR, McIntosh JM, Jones RM;
PI WPI: 2000-505965/45.
DR N-PSDB; AAA89455.
XX alpha-conotoxin polypeptides derived from the venom of cone snails
PT useful e.g. as neuromuscular blocking agents for use in surgery and for
PT treating unipolar depression -
XX Claim 39; Page 47; 229pp; English.

XX The present invention relates to a number of alpha-conotoxin peptides and
CC their coding sequences from a number of different species of cone snail.
CC These peptides are found in minute quantities in the venom of the snails,
CC and are targeted at the neuronal nicotinic acetylcholine receptors of the
CC nervous system. They usually contain two disulphide bonds, which give
CC them defined conformations, a rarity in molecules this small. The
CC alpha-conotoxins can be used as neuromuscular blocking agents in surgery,
CC acetylcholine receptors, including cardiovascular disorders, gastric
CC motility disorders, urinary incontinence, nicotine addiction, mood
CC disorders such as bipolar disorder, unipolar depression, dysthymia and
CC seasonal affective disorder, and small cell lung carcinoma.
XX Sequence 35 AA;

Query Match 100.0%; Score 49; DB 21; Length 35;
Best Local Similarity 23.5%; Pred. No. 6.2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
Db 19 iawdicccpdcnhkev 35

RESULT 50
ABG29284
ID ABG29284 standard; Protein; 35 AA.
XX AC ABG29284;
XX AC ABG29284;
DT 13-FEB-2002 (first entry)
DE Novel human diagnostic protein #29275.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS93471.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

PS Claim 20; SEQ ID No 59643; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 35 AA;

Query Match 100.0%; Score 49; DB 22; Length 35;

Best Local Similarity 23.5%; Pred. No. 6.2e+02;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Oy 1 XXXXXCXXCXXCXXCX 17

Db 1 KSKSCCCCCCCCCCCC 17

Search completed: July 1, 2002, 12:27:39
Job time: 161 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 12:25:23 ; Search time 12.96 Seconds
(without alignments)
32.040 Million cell updates/sec

Title: US-09-493-795A-1

Perfect score: 49

Sequence: 1 XXXXXCXXXXXXCXXC 17

Scoring table:

BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database :

Issued Patents_AA:*

1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*

2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*

3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*

4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*

5: /cgn2_6/ptodata/2/1aa/PCTus_COMB.pep:*

6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 49 | 100.0 | 26 | 4 | US-09-227-357-579 |
| 2 | 49 | 100.0 | 29 | 1 | US-08-428-248-6 |
| 3 | 49 | 100.0 | 29 | 1 | US-08-319-554A-1 |
| 4 | 49 | 100.0 | 29 | 1 | US-08-624-123-1 |
| 5 | 49 | 100.0 | 29 | 1 | US-08-775-476-1 |
| 6 | 49 | 100.0 | 29 | 1 | US-09-018-799-1 |
| 7 | 49 | 100.0 | 29 | 2 | US-09-099-974-1 |
| 8 | 49 | 100.0 | 29 | 2 | US-09-005-156-1 |
| 9 | 49 | 100.0 | 29 | 4 | US-09-136-769A-5 |
| 10 | 49 | 100.0 | 29 | 4 | US-09-136-769A-16 |
| 11 | 49 | 100.0 | 30 | 1 | US-08-428-248-1 |
| 12 | 49 | 100.0 | 30 | 1 | US-08-428-248-8 |
| 13 | 49 | 100.0 | 31 | 1 | US-08-428-248-5 |
| 14 | 49 | 100.0 | 31 | 1 | US-08-428-248-9 |
| 15 | 49 | 100.0 | 31 | 1 | US-08-775-477-1 |
| 16 | 49 | 100.0 | 31 | 1 | US-08-775-476-2 |
| 17 | 49 | 100.0 | 31 | 1 | US-09-018-799-2 |
| 18 | 49 | 100.0 | 31 | 2 | US-09-099-974-2 |
| 19 | 49 | 100.0 | 31 | 2 | US-09-010-204-1 |
| 20 | 49 | 100.0 | 38 | 1 | US-07-609-716-53 |
| 21 | 49 | 100.0 | 38 | 2 | US-08-902-516-47 |
| 22 | 49 | 100.0 | 41 | 5 | PCT-US96-01720-7 |
| 23 | 49 | 100.0 | 45 | 4 | US-08-900-230-7 |
| 24 | 49 | 100.0 | 45 | 4 | US-08-900-230-10 |
| 25 | 49 | 100.0 | 45 | 4 | US-08-900-230-18 |
| 26 | 49 | 100.0 | 45 | 4 | US-08-900-230-52 |
| 27 | 49 | 100.0 | 46 | 4 | US-08-900-230-40 |

| | | | | | | |
|----|----|-------|-----|---|-------------------|--------------------|
| 28 | 49 | 100.0 | 47 | 3 | US-08-482-085B-91 | Sequence 91, Appl |
| 29 | 49 | 100.0 | 50 | 4 | US-08-900-230-8 | Sequence 8, Appl |
| 30 | 49 | 100.0 | 53 | 1 | US-08-239-256-2 | Sequence 2, Appl |
| 31 | 49 | 100.0 | 53 | 4 | US-09-305-639-2 | Sequence 2, Appl |
| 32 | 49 | 100.0 | 55 | 3 | US-08-476-509B-46 | Sequence 46, Appl |
| 33 | 49 | 100.0 | 57 | 1 | US-07-609-716-56 | Sequence 56, Appl |
| 34 | 49 | 100.0 | 62 | 1 | US-08-137-800-51 | Sequence 51, Appl |
| 35 | 49 | 100.0 | 62 | 1 | US-08-477-383-51 | Sequence 51, Appl |
| 36 | 49 | 100.0 | 62 | 1 | US-08-487-174-51 | Sequence 51, Appl |
| 37 | 49 | 100.0 | 62 | 1 | US-08-480-750-51 | Sequence 51, Appl |
| 38 | 49 | 100.0 | 68 | 1 | US-07-689-693B-11 | Sequence 11, Appl |
| 39 | 49 | 100.0 | 84 | 4 | US-09-386-493-8 | Sequence 8, Appl |
| 40 | 49 | 100.0 | 84 | 4 | US-09-386-493-14 | Sequence 14, Appl |
| 41 | 49 | 100.0 | 109 | 2 | US-08-527-044-2 | Sequence 2, Appl |
| 42 | 49 | 100.0 | 109 | 3 | US-09-013-780-2 | Sequence 2, Appl |
| 43 | 49 | 100.0 | 120 | 3 | US-08-508-761B-22 | Sequence 22, Appl |
| 44 | 49 | 100.0 | 131 | 6 | 5177193-3 | Patent No. 5177193 |
| 45 | 49 | 100.0 | 197 | 2 | US-08-855-261A-1 | Sequence 1, Appl |
| 46 | 49 | 100.0 | 197 | 4 | US-09-227-224-1 | Sequence 1, Appl |
| 47 | 49 | 100.0 | 198 | 4 | US-09-227-357-232 | Sequence 232, App |
| 48 | 49 | 100.0 | 202 | 1 | US-08-048-700-2 | Sequence 2, Appl |
| 49 | 49 | 100.0 | 202 | 2 | US-08-855-261A-3 | Sequence 3, Appl |
| 50 | 49 | 100.0 | 202 | 3 | US-08-839-711-3 | Sequence 3, Appl |

ALIGNMENTS

RESULT 1
US-09-227-357-579
; Sequence 579, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722

;; EARLIER FILING DATE: 1997-08-18
;; EARLIER APPLICATION NUMBER: 60/055,723
;; EARLIER FILING DATE: 1997-08-18
;; EARLIER APPLICATION NUMBER: 60/055,948
;; EARLIER FILING DATE: 1997-08-18
;; EARLIER APPLICATION NUMBER: 60/055,949
;; EARLIER FILING DATE: 1997-08-18
;; EARLIER APPLICATION NUMBER: 60/055,953
;; EARLIER FILING DATE: 1997-08-18
;; EARLIER APPLICATION NUMBER: 60/055,950
;; EARLIER FILING DATE: 1997-08-18
;; EARLIER APPLICATION NUMBER: 60/055,947
;; EARLIER FILING DATE: 1997-08-18
;; EARLIER APPLICATION NUMBER: 60/055,964
;; EARLIER FILING DATE: 1997-08-18
;; EARLIER APPLICATION NUMBER: 60/056,360
;; EARLIER FILING DATE: 1997-08-18
;; EARLIER APPLICATION NUMBER: 60/055,684
;; EARLIER FILING DATE: 1997-08-18
;; EARLIER APPLICATION NUMBER: 60/055,984
;; EARLIER FILING DATE: 1997-08-18
;; EARLIER APPLICATION NUMBER: 60/055,954
;; EARLIER FILING DATE: 1997-08-18
;; EARLIER APPLICATION NUMBER: 60/058,785
;; EARLIER FILING DATE: 1997-09-12
;; EARLIER APPLICATION NUMBER: 60/058,664
;; EARLIER FILING DATE: 1997-09-12
;; EARLIER APPLICATION NUMBER: 60/058,660
;; EARLIER FILING DATE: 1997-09-12
;; EARLIER APPLICATION NUMBER: 60/058,661
;; EARLIER FILING DATE: 1997-09-12
;; NUMBER OF SEQ ID NOS: 672
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 579
;; LENGTH: 26
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-227-357-579

Query Match 100.0%; Score 49; DB 4; Length 26;
Best Local Similarity 23.5%; Pred. No. 1.3e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCXXXXXXCX 17
Db 5 GTSNCCRGCCGSVCI 21

RESULT 2
US-08-428-248-6
; Sequence 6, Application US/08428248
; Patent No. 5627154
; GENERAL INFORMATION:
; APPLICANT: Kelbaugh, Paul R.
; APPLICANT: Saccomano, Nicholas A.
; APPLICANT: Volkman, Robert A.
; TITLE OF INVENTION: CALCIUM CHANNEL BLOCKING POLYPEPTIDES
; TITLE OF INVENTION: FROM HETEROPODA VENATORIA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pfizer Inc
; STREET: 235 East 42nd Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/428,248
;; FILING DATE: 27-APRIL-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/970333
;; FILING DATE: 30-OCTOBER-1992
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 29 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; ORIGINAL SOURCE:
;; ORGANISM: Heteropoda venatoria
;; TISSUE TYPE: venom
US-08-428-248-6

Query Match 100.0%; Score 49; DB 1; Length 29;
Best Local Similarity 23.5%; Pred. No. 1.4e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCXXXXXXCX 17
Db 11 DTSKDCGYVCHLWCK 27

RESULT 3
US-08-319-554A-1
; Sequence 1, Application US/08319554A
; Patent No. 5719264
; GENERAL INFORMATION:
; APPLICANT: Shon, Ki-Joon
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Hasson, Arik
; APPLICANT: Spira, Micha E.
; TITLE OF INVENTION: Conotoxin Peptides
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue N.W., Suite 1000
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/319,554A
; FILING DATE: 07-OCT-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24260-107674-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Conus gloriamaris
; STRAIN: GmVIA


```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/099,974
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/018,799
; FILING DATE:
; APPLICATION NUMBER: US/08/775,476
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hohenschutz, Liza D.
; REGISTRATION NUMBER: 33,712
; REFERENCE/DOCKET NUMBER: PHM.70122
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 886-7466
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 29
; OTHER INFORMATION: /note= "xaa is amidated leucine"
;
US-09-099-974-1

Query Match 100.0%; Score 49; DB 2; Length 29;
Best Local Similarity 23.5%; Pred.No.1.4e+02;
Matches 4; Conservative 13; Mismatches 0; Indels

QY 1 XXXXXCCXXXCXXCX 17
Db 10 DSERKCCEDMWCLNCK 26
:::|||||:::|:|:|

RESULT 8
US-09-006-156-1
; Sequence 1, Application US/09006156
; Patent No. 5950295
; GENERAL INFORMATION:
; APPLICANT: Shon, Ki-Joon
; APPLICANT: Agintova, Baldomero M.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Hasson, Arik
; APPLICANT: Spira, Micha E.
; TITLE OF INVENTION: Conotoxin Peptides
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue N.W., Suite 1000
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/006,156
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/319,554
; FILING DATE: 07-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrev L.

```


REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24260-107674-1
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Conus gloriamaris
STRAIN: GmVIA
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 4..19
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 11..24
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 18..28
US-09-006-156-1

Query Match 100.0%; Score 49; DB 2; Length 29;
Best Local Similarity 23.5%; Pred. No. 1.4e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 13 PIFQNCRCGNCVLCV 29

RESULT 9
US-09-136-769A-5
Sequence 5, Application US/09136769A
Patent No. 6307014
GENERAL INFORMATION:
APPLICANT: Furie, Bruce
APPLICANT: Furie, Barbara
APPLICANT: Stenflo, Johan
APPLICANT: Rigby, Alan C.
APPLICANT: Roepstoft, Peter
TITLE OF INVENTION: CONOPEPTIDES
FILE REFERENCE: 50065/002001
CURRENT APPLICATION NUMBER: US/09/136,769A
CURRENT FILING DATE: 1998-08-19
NUMBER OF SEQ ID NOS: 26
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 29
TYPE: PRT
ORGANISM: Conus textile
FEATURE:
NAME/KEY: VARIANT
LOCATION: (4)...(4)
OTHER INFORMATION: Xaa is gamma-carboxyglutamic acid
US-09-136-769A-5

Query Match 100.0%; Score 49; DB 4; Length 29;
Best Local Similarity 23.5%; Pred. No. 1.4e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 9 SSSGSCCHKSCCRWTCN 25

RESULT 10
US-09-136-769A-16
Sequence 16, Application US/09136769A
Patent No. 6307014
GENERAL INFORMATION:
APPLICANT: Furie, Bruce
APPLICANT: Furie, Barbara
APPLICANT: Stenflo, Johan
APPLICANT: Rigby, Alan C.
APPLICANT: Roepstoft, Peter
TITLE OF INVENTION: CONOPEPTIDES
FILE REFERENCE: 50065/002001
CURRENT APPLICATION NUMBER: US/09/136,769A
CURRENT FILING DATE: 1998-08-19
NUMBER OF SEQ ID NOS: 26
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16
LENGTH: 29
TYPE: PRT
ORGANISM: Conus textile
FEATURE:
NAME/KEY: VARIANT
LOCATION: (4)...(4)
OTHER INFORMATION: Xaa is gamma-carboxyglutamic acid.
US-09-136-769A-16

Query Match 100.0%; Score 49; DB 4; Length 29;
Best Local Similarity 23.5%; Pred. No. 1.4e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 9 SSSGSCCHKSCCRWTCN 25

RESULT 11
US-08-428-248-1
Sequence 1, Application US/08428248
Patent No. 5627154
GENERAL INFORMATION:
APPLICANT: Kelbaugh, Paul R.
APPLICANT: Saccamano, Nicholas A.
APPLICANT: Volkmann, Robert A.
TITLE OF INVENTION: CALCIUM CHANNEL BLOCKING POLYPEPTIDES
TITLE OF INVENTION: FROM HETEROPODA VENATORIA
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pfizer Inc
STREET: 235 East 42nd Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10017
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/428,248
FILING DATE: 27-APRIL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/970333
FILING DATE: 30-OCTOBER-1992
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide

; ORIGINAL SOURCE:
 ; ORGANISM: Heteropoda venatoria
 ; TISSUE TYPE: venom
 US-08-428-248-1

Query Match 100.0%; Score 49; DB 1; Length 30;
Best Local Similarity 23.5%; Pred. NO. 1.4e+02;
Matches 4; Conservative 13; Mismatches 0; Indels

```
Qy 1 XXXXXCCXXXCXXCXX 17
    ::::|:::|:::|:::|:
Db 11 DSNADCCGYYVCRWLCK 27
```

RESULT 12
US-08-428-248-8
; Sequence 8, Application US/08428248
; Patent No. 5627154
; GENERAL INFORMATION:
; APPLICANT: Kelbaugh, Paul R.
; APPLICANT: Saccomano, Nicholas A.
; APPLICANT: Volkmann, Robert A.
; TITLE OF INVENTION: CALCIUM CHANNEL BLOCKING POLYPEPTIDES
; TITLE OF INVENTION: FROM HETEROPODA VENATORIA

ORDER OF DISBURSEMENT:
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pfizer Inc
STREET: 235 East 42nd Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10017

```

;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: PatentIn Release #1.0,
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: US/08/428,248
;
; FILING DATE: 27-APRIL-1995
;

```

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/970333
 ; FILING DATE: 30-OCTOBER-1992
 ; INFORMATION FOR SEQ ID NO: 8:

LENGTH: 30 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Heteropoda venatoria
 TISSUE TYPE: venom
 US-08-428-248-8

```
Query Match      100.0%; Score 49; DB 1; Length 30;
Best Local Similarity 23.5%; Pred. No. 1.4e+02;
Matches 4; Conservative 13; Mismatches 0; Indels
```

```
Qy      1 XXXXXCCXXXCXXCXXCX 17
        ::::|:::|:::|:::|:
Db     11 DTNADCCGYVCRLWCK 27
```

RESULT 13
US-08-428-248-5
; Sequence 5, Application US/08428248
; Patent No. 5637154
; GENERAL INFORMATION:
; APPLICANT: Kelbaugh, Paul R.
; APPLICANT: Saccomano, Nicholas A.

```

; APPLICANT: Volkman, Robert A.
; TITLE OF INVENTION: CALCIUM CHANNEL BLOCKING POLYPEPTIDES
; TITLE OF INVENTION: FROM HETEROPODA VENATORIA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pfizer Inc
; STREET: 235 East 42nd Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10017

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/428,248
FILING DATE: 27-APRIL-1995

```

; PRIOR APPLICATION DATA: US 07/970333
 ; APPLICATION NUMBER: US 07/970333
 ; FILING DATE: 30-OCTOBER-1992
 ;

```

; INFORMATION FOR SEQ ID NO: 5:
;
; SEQUENCE CHARACTERISTICS:

```

```

LENGTH: 31 amino acids
TYPE: amino acid

```

```

; STRANDEDNESS: single
; TOPOLOGY: linear

```

MOLECULE TYPE: peptide
ORIGINAL SOURCE:

ORGANISM: Heteropod

US-08-428-248-5
TISSUE TYPE: venom
;

Query Match 100.0%; Score 49; DB 1; Length 31;
Best Local Similarity 23.5%; pred. No. 1.4e+02;
Matches 4; Conservative 13; Mismatches 0; Indels

```
QY 1 XXXXXCCXXXCXXCXX 17
    : : : : : : : : : : : :
Db 10 STHADCCGEFICKLWCR 26
```

RESULT 14
US-08-428-248-9
; Sequence 9, Application US/08428248

Patent No. 5627154

APPLICANT: Kelbaugh, Paul R.

APPLICANT: SACCOMANO, NICHOLAS R.
APPLICANT: VOLKMAN, ROBERT A.

;;
; TITLE OF INVENTION: CALCIUM CHANNEL BLOCKING POLYPEPTIDES
; TITLE OF INVENTION: FROM HETEROPODA VENATORIA

```

; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:

```

ADDRESSEE: Pfizer Inc
STREET: 235 East 42nd Street

CITY: New York

STATE: NEW YORK
COUNTRY: U.S.A.

```

; ZIP: 10017
; COMPUTER READABLE FORM:

```

```

;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
COMPUTER: Datacube Datacube #10

```

; SOFTWARE: Patent In Release #1.0,
;
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/428,248
FILING DATE: 27-APRIL-1995

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/970333

FILING DATE: 30-OCTOBER-1992

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/775,476
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hohenschutz, Liza D.
; REGISTRATION NUMBER: 33,712
; REFERENCE/DOCKET NUMBER: PHM.70122
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 886-7466
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-018-799-2

Query Match 100.0%; Score 49; DB 1; Length 31;
Best Local Similarity 23.5%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCXXXXXXCX 17
Db 10 DEERKCEGLVCLWCK 26

RESULT 18
US-09-099-974-2
; Sequence 2, Application US/09099974
; Patent No. 5877026
; GENERAL INFORMATION:
; APPLICANT: Lampe, Richard A.
; TITLE OF INVENTION: Analgesic Peptides from Venom of
; TITLE OF INVENTION: Grammostola Spatulata and Use Thereof
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZENECA Pharmaceuticals
; STREET: 1800 Concord Pike
; CITY: Wilmington
; STATE: DE
; COUNTRY: United States
; ZIP: 19850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/099,974
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/018,799
; FILING DATE:
; APPLICATION NUMBER: US/08/775,476
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hohenschutz, Liza D.
; REGISTRATION NUMBER: 33,712
; REFERENCE/DOCKET NUMBER: PHM.70122
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 886-7466
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-099-974-2

Query Match 100.0%; Score 49; DB 2; Length 31;

Best Local Similarity 23.5%; Pred. No. 1.4e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCXXXXXXCX 17
Db 10 DEERKCEGLVCLWCK 26

RESULT 19
US-09-010-204-1
; Sequence 1, Application US/09010204
; Patent No. 5968838
; GENERAL INFORMATION:
; APPLICANT: Lampe, Richard A.
; TITLE OF INVENTION: Antiarhythmic Peptide From Venom of
; TITLE OF INVENTION: Spider Grammostola spatulata
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZENECA Pharmaceuticals
; STREET: 1800 Concord Pike
; CITY: Wilmington
; STATE: DE
; COUNTRY: United States
; ZIP: 19850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/010,204
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/775,477
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hohenschutz, Liza D.
; REGISTRATION NUMBER: 33,712
; REFERENCE/DOCKET NUMBER: PHM.70123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 886-7466
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-010-204-1

Query Match 100.0%; Score 49; DB 2; Length 31;
Best Local Similarity 23.5%; Pred. No. 1.4e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCXXXXXXCX 17
Db 10 DEERKCEGLVCLWCK 26

RESULT 20
US-07-609-716-53
; Sequence 53 Application US/07609716
; Patent No. 5514581
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; TITLE OF INVENTION: Functional Recombinantly Prepared
; TITLE OF INVENTION: Synthetic Protein Polymer
; NUMBER OF SEQUENCES: 118
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert

STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/609,716
FILING DATE: 06-NOV-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: 20015
REFERENCE/DOCKET NUMBER: A-55186-3/BIR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-07-609-716-53

Query Match 100.0%; Score 49; DB 1; Length 38;
Best Local Similarity 23.5%; Pred. No. 1.7e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCXXXXXXCXCX 17
Db 5 GTCCTCCAGGACCGCA 21

RESULT 21
US-08-902-516-47
Sequence 47, Application US/08902516
Patent No. 5891432
GENERAL INFORMATION:
APPLICANT: SOO HOO, William
TITLE OF INVENTION: MEMBRANE-BOUND CYTOKINE COMPOSITIONS
TITLE OF INVENTION: COMPRISING GM-CSF AND METHODS OF MODULATING AN IMMUNE
TITLE OF INVENTION: RESPONSE USING SAME
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: CAMPBELL & FLORES, LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92121
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/902,516
FILING DATE: 29-JUL-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-1M 2442
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)535-9001
TELEFAX: (619)535-8949

INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-902-516-47

Query Match 100.0%; Score 49; DB 2; Length 38;
Best Local Similarity 23.5%; Pred. No. 1.7e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCXXXXXXCXCX 17
Db 5 AGGGCCCTAGCACCCA 21

RESULT 22
PCT-US96-01720-7
Sequence 7, Application PC/TUS9601720
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: MODIFIED-AFFINITY STREPTAVIDIN
NUMBER OF SEQUENCES: 11
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01720
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/387,055
FILING DATE: 09-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 16336-5PC
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-01720-7

Query Match 100.0%; Score 49; DB 5; Length 41;
Best Local Similarity 23.5%; Pred. No. 1.8e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCXXXXXXCXCX 17
Db 11 GACCTCCGACCAACCG 27

RESULT 23
US-08-900-230-7
Sequence 7, Application US/08900230
Patent No. 6329197
GENERAL INFORMATION:
APPLICANT: Bard, Jonathan A.
TITLE OF INVENTION: DNA ENCODING GALANIN GALR3 RECEPTORS AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of The Americas
CITY: New York

; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/900,230
; FILING DATE: 23-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 52241-C/JPW/ADM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; MOLECULE TYPE:
; TOPOLOGY: linear
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-900-230-7

Query Match 100.0%; Score 49; DB 4; Length 45;
Best Local Similarity 23.5%; Pred. No. 1.9e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCCXXXCCXXXCX 17
Db 3 CACGCCAGCACCAGCG 19

RESULT 24
US-08-900-230-10
; Sequence 10, Application US/08900230
; Patent No. 6329197
; GENERAL INFORMATION:
; APPLICANT: Bard, Jonathan A.
; TITLE OF INVENTION: DNA ENCODING GALANN GALR3 RECEPTORS AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of The Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/900,230
; FILING DATE: 23-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 52241-C/JPW/ADM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-900-230-10

Query Match 100.0%; Score 49; DB 4; Length 45;
Best Local Similarity 23.5%; Pred. No. 1.9e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCCXXXCCXXXCX 17
Db 14 CATTTCCTCATCTTCT 30

RESULT 25
US-08-900-230-18
; Sequence 18, Application US/08900230
; Patent No. 6329197
; GENERAL INFORMATION:
; APPLICANT: Bard, Jonathan A.
; TITLE OF INVENTION: DNA ENCODING GALANN GALR3 RECEPTORS AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of The Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/900,230
; FILING DATE: 23-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 52241-C/JPW/ADM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; MOLECULE TYPE:
; TOPOLOGY: linear
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-900-230-18

Query Match 100.0%; Score 49; DB 4; Length 45;
Best Local Similarity 23.5%; Pred. No. 1.9e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCCXXXCCXXXCX 17
Db 22 AGTTTCCTATGCCAACT 38

```

RESULT 27
US-08-900-230-40
; Sequence 40, Application US/08900230
; Patent No. 6329197
; GENERAL INFORMATION:
; APPLICANT: Bard, Jonathan A.
; TITLE OF INVENTION: DNA ENCODING GALANIN GALR3 RECEPTORS AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of The Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; GENERAL INFORMATION:
;
; APPLICANT: Bard, Jonathan A.
;
; TITLE OF INVENTION: DNA ENCODING GALANN GALR3 RECEPTORS AND
;
; TITLE OF INVENTION: USES THEREOF
;
; NUMBER OF SEQUENCES: 59
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: Cooper & Dunham LLP
;
; STREET: 1185 Avenue of The Americas
;
; CITY: New York
;
; STATE: New York
;
; COUNTRY: U.S.A.
;
; ZIP: 11036
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/175,155
;; FILING DATE: 29-DEC-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Trecartin, Richard F.
;; REGISTRATION NUMBER: 31,801
;; REFERENCE/DOCKET NUMBER: A-55186-6/RFT/MTK
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-781-1989
;; TELEFAX: 415-398-3249
;; INFORMATION FOR SEQ ID NO: 91:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 47 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-08-482-085B-91

Query Match 100.0%; Score 49; DB 3; Length 47;
Best Local Similarity 23.5%; Pred. No. 2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
Db 25 ATGCACCATCTCGAGCA 41

RESULT 29
US-08-900-230-8
; Sequence 8, Application US/08900230
; Patent No. 6329197
; GENERAL INFORMATION:
; APPLICANT: Bard, Jonathan A.
; TITLE OF INVENTION: DNA ENCODING GALANN GALR3 RECEPTORS AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of The Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/900,230
; FILING DATE: 23-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 52241-C/JPW/ADM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-900-230-8

Query Match 100.0%; Score 49; DB 4; Length 50;
Best Local Similarity 23.5%; Pred. No. 2.1e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
Db 9 ACCAACCTGTTTCATCCT 25

RESULT 30
US-08-239-256-2
; Sequence 2, Application US/08239256
; Patent No. 5585345
; GENERAL INFORMATION:
; APPLICANT: BOIME, IRVING
; APPLICANT: MATZUK, MARTIN M.
; APPLICANT: KEENE, JEFFREY L.
; TITLE OF INVENTION: CTP EXTENDED FORM OF LH
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/239,256
; FILING DATE: 06-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 29500-20030.12
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 53 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-239-256-2

Query Match 100.0%; Score 49; DB 1; Length 53;
Best Local Similarity 23.5%; Pred. No. 2.2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
Db 6 FFFLCCWKAICNSCE 22

RESULT 31
US-09-305-639-2
; Sequence 2, Application US/09305639
; Patent No. 6200778
; GENERAL INFORMATION:
; APPLICANT: Treco, Douglas A.
; APPLICANT: Heartlein, Michael W.
; APPLICANT: Selden, Richard F.
; TITLE OF INVENTION: GENOMIC SEQUENCES FOR PROTEIN PRODUCTION AND DELIVERY
; FILE REFERENCE: 07236/016001
; CURRENT APPLICATION NUMBER: US/09/305,639
; CURRENT FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: 60/084,663

EARLIER FILING DATE: 1998-05-07
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 53
TYPE: PRT
ORGANISM: Homo sapiens
US-09-305-639-2

Query Match 100.0%; Score 49; DB 4; Length 53;
Best Local Similarity 23.5%; Pred. No. 2.2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCXXXXXXCXXCX 17

Db 6 FFFLCWCRAICNCSE 22

RESULT 32
US-08-476-509B-46
Sequence 45, Application US/08476509B
Patent No. 6034212
GENERAL INFORMATION:
APPLICANT: SUDOL, MARIUS
APPLICANT: PEER, BORK
APPLICANT: HENRY, CHEN
TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A
TITLE OF INVENTION: SIGNALING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE
TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC USES
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauder & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,509B
FILING DATE: 01-DEC-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-101 CIP
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 55 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "PRIMER"
HYPOTHETICAL: NO
US-08-476-509B-46

Query Match 100.0%; Score 49; DB 3; Length 55;
Best Local Similarity 23.5%; Pred. No. 2.3e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCXXXXXXCXXCX 17

Db 14 GCACACCGCCACCTCCT 30

RESULT 33
US-07-609-716-56
Sequence 56, Application US/07609716
Patent No. 5514581
GENERAL INFORMATION:
APPLICANT: Ferrari, Franco A.
APPLICANT: Cappello, Joseph
TITLE OF INVENTION: Functional Recombinantly Prepared
TITLE OF INVENTION: Synthetic Protein Polymer
NUMBER OF SEQUENCES: 118
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/609,716
FILING DATE: 06-NOV-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: 20015
REFERENCE/DOCKET NUMBER: A-55186-3/BIR
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-07-609-716-56

Query Match 100.0%; Score 49; DB 1; Length 57;
Best Local Similarity 23.5%; Pred. No. 2.3e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCXXXXXXCXXCX 17

Db 23 TGGCAGCGGGCTACCT 39

RESULT 34
US-08-137-800-51
Sequence 51, Application US/08137800
Patent No. 5514774
GENERAL INFORMATION:
APPLICANT: Olivera, Baldomero M.
APPLICANT: Cruz, Lourdes J.
APPLICANT: Hilliard, David R.
APPLICANT: McIntosh, J. Michael
APPLICANT: Santos, Ameurfin D.
TITLE OF INVENTION: Conotoxin Peptides
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti
STREET: 1201 New York Avenue N.W., Suite 1000
CITY: Washington
STATE: DC

Query Match 100.0%; Score 49; DB 1; Length 62;
Best Local Similarity 23.5%; Pred. No. 2.5e+02;
Matches 4; Conservative 13; Mismatches 0; Indels

```
QY      1 XXXXXCCXXXCCXXCX 17
        ::::|:::|:::|:::|:
Db     46 FIRDYCCHRGPCMVWCG 62
```

RESULT 36
US-08-487-174-51
; Sequence 51, Application US/08487174
; Patent No. 5595972

```

: GENERAL INFORMATION:
: APPLICANT: Olivera, Baldomero M.
: APPLICANT: Cruz, Lourdes J.
: APPLICANT: Hillyard, David R.
: APPLICANT: Macintosh, J. Michael
: APPLICANT: Santos, Ameurino S.
: TITLE OF INVENTION: Conotoxin Peptides
: NUMBER OF SEQUENCES: 59
: CORRESPONDENCE ADDRESS:
:

```

APPLICANT: SANTOS, Ameloxin S.
 TITLE OF INVENTION: Conotoxin Peptides
 NUMBER OF SEQUENCES: 59
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Venable, Baetjer, Howard & Civiletti
 STREET: 1201 New York Avenue, N.W., Suite 1000
 CITY: Washington
 STATE: DC
 COUNTRY: U.S.A.
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/487,174
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/137,800
 FILING DATE: 19-OCT-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/084,848
 FILING DATE: 29-JUN-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Ihnen, Jeffrey L.
 REGISTRATION NUMBER: 28,957
 REFERENCE/DOCKET NUMBER: 24260-107673
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-962-4810
 TELEFAX: 202-962-8300
 INFORMATION FOR SEQ ID NO: 51:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 62 amino acids

TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Conus bandanus
US-08-487-174-51

Query Match 100.0%; Score 49; DB 1; Length 62;
Best Local Similarity 23.5%; Pred. No. 2.5e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCXXXXXXCX 17
Db 46 FIRDYCHRGPCVMWCG 62

RESULT 37
US-08-480-750-51
; Sequence 51, Application US/08480750
; Patent No. 563347
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Hillyard, David R.
; APPLICANT: Macintosh, J. Michael
; APPLICANT: Santos, Ameurfin S.
; TITLE OF INVENTION: Conotoxin Peptides
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,750
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/137,800
; FILING DATE: 19-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/084,848
; FILING DATE: 29-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24260-107673
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 62 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Conus bandanus
US-08-480-750-51

Query Match 100.0%; Score 49; DB 1; Length 62;
Best Local Similarity 23.5%; Pred. No. 2.5e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCXXXXXXCX 17
Db 46 FIRDYCHRGPCVMWCG 62

RESULT 38
US-07-689-693B-11
; Sequence 11, Application US/07689693B
; Patent No. 5231011
; GENERAL INFORMATION:
; APPLICANT: David Hillyard
; APPLICANT: Baldomero M. Olivera
; TITLE OF INVENTION: Segregated Folding Determinants
; TITLE OF INVENTION: for Small Disulfide-Rich Peptides
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thorpe, No. 5231011th & Western
; STREET: 9035 South 700 East, Suite 200
; CITY: Sandy
; STATE: Utah
; COUNTRY: USA
; ZIP: 84070
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 720 Kb storage
; COMPUTER: Compaq LTE/286
; OPERATING SYSTEM: DOS 4.01
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/689,693B
; FILING DATE: 19910418
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: none
; FILING DATE: na
; ATTORNEY/AGENT INFORMATION:
; NAME: Western, M. Wayne
; REGISTRATION NUMBER: 22,788
; REFERENCE/DOCKET NUMBER: 9925
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (801) 566-6633
; TELEFAX: (801) 566-0750
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Prepropeptide sequence for three loop
; NAME/KEY: conotoxin from Cirlcer conotoxin from Conus
; NAME/KEY: textile
; IDENTIFICATION METHOD: Libraries were created
; IDENTIFICATION METHOD: using oligo-dT primed puc13 vector
US-07-689-693B-11

Query Match 100.0%; Score 49; DB 1; Length 68;
Best Local Similarity 23.5%; Pred. No. 2.7e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCXXXXXXCX 17
Db 48 DPNKRCPPVACNMCK 64

RESULT 39
US-09-386-493-8
; Sequence 8, Application US/09386493
; Patent No. 6262247

```

; TITLE OF INVENTION: Coccidiosis poultry vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AKZO NO. 5885568el N.V
; STREET: 1300 PICCARD DRIVE #206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850-4373
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/527,044
; FILING DATE: 12-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 94202667.6
; FILING DATE: 16-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: GORMLEY, MARY E.
; REGISTRATION NUMBER: 34,409
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-527-044-2

Query Match 100.0%; Score 49; DB 2; Length 109;
Best Local Similarity 23.5%; Pred No. 3.9e+02;
Matches 4; Conservative 13; Mismatches 0; Indels

QY 1 XXXXXCCXXXCCXXXCC 17
; : : : : : : : : : : : :
Db 14 RYTNCCGSKYCCSKCC 30

RESULT 42
US-09-013-780-2
; Sequence 2, Application US/09013780
; Patent No. 6001363
; GENERAL INFORMATION:
; APPLICANT: BUNSTEAD, Janene Marilyn J.M.
; APPLICANT: TOMLEY, Fiona Margaret F.
; APPLICANT: DUNN, Patrick Paul James P.
; APPLICANT: VERUEULEN, Arnoldus Nicolaas A.N.
; TITLE OF INVENTION: Coccidiosis poultry vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AKZO NO. 6001363el N.V
; STREET: 1300 PICCARD DRIVE #206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850-4373
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/013,780
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/527,044
; FILING DATE: 12-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: GORMLEY, MARY E.
; REGISTRATION NUMBER: 34,409

```


; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: COLNNOT19
; CLONE: 1634851
US-08-855-261A-1

Query Match 100.0%; Score 49; DB 2; Length 197;
Best Local Similarity 23.5%; Pred. No. 6.2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Oy 1 XXXXXCXXXXCXXCX 17
Db 70 AGKGCCGCGCCGNCR 86

RESULT 46
US-09-227-224-1
; Sequence 1, Application US/09227224
; Patent No. 6350581
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: NEW TUMOR-ASSOCIATED ANTIGEN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/227,224
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/855,261
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0298 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 197 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

IMMEDIATE SOURCE:
; LIBRARY: COLNNOT19
; CLONE: 1634851
US-09-227-224-1

Query Match 100.0%; Score 49; DB 4; Length 197;
Best Local Similarity 23.5%; Pred. No. 6.2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Oy 1 XXXXXCXXXXCXXCX 17
Db 70 AGKGCCGCGCCGNCR 86

RESULT 47
US-09-227-357-232
; Sequence 232, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,660

EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
EARLIER FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 672
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 232
LENGTH: 198
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (198)
OTHER INFORMATION: Xaa equals stop translation
US-09-227-357-232

Query Match 100.0%; Score 49; DB 4; Length 198;
Best Local Similarity 23.5%; Pred. No. 6.2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
Db 70 AGGCGCGAGCGGNCR 86

RESULT 48

US-08-048-700-2
Sequence 2, Application US/08048700
Patent No. 5597707

GENERAL INFORMATION:
APPLICANT: Marken, John
APPLICANT: Schieven, Gary L.
APPLICANT: Hellstrom, Ingegerd
APPLICANT: Hellstrom, Karl E.
APPLICANT: Aruffo, Alejandro
TITLE OF INVENTION: A Glycoprotein Tumor Associated Antigen
TITLE OF INVENTION: Recognized By The Monoclonal Antibody L6, Its
TITLE OF INVENTION: Oligonucleotide Sequence and Methods For Their Use
Patent No. 5597707
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bristol-Myers Squibb Company
STREET: 3005 First Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98121

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/048,700
FILING DATE: 15-APR-1993
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
NAME: Poor, Brian W.
REGISTRATION NUMBER: 32,928
REFERENCE/DOCKET NUMBER: ON0089-
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)728-4800
TELEFAX: (206)727-3601

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 202 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Colon

CELL TYPE: carcinoma
CELL LINE: H3347
US-08-048-700-2

Query Match 100.0%; Score 49; DB 1; Length 202;
Best Local Similarity 23.5%; Pred. No. 6.3e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
Db 73 DCCGCCGHCNCGRCA 89

RESULT 49

US-08-855-261A-3
Sequence 3, Application US/08855261A
Patent No. 5922566

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goli, Surya K.
APPLICANT: Shah, Purvi
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: NEW TUMOR-ASSOCIATED ANTIGEN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/855,261A
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0298 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 202 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 186804

US-08-855-261A-3

Query Match 100.0%; Score 49; DB 2; Length 202;
Best Local Similarity 23.5%; Pred. No. 6.3e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
Db 73 DCCGCCGHCNCGRCA 89

RESULT 50

US-08-839-711-3

Tue Jul 2 09:11:47 2002

; Sequence 3, Application US/08839711
; Patent No. 6033870
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: NOVEL HUMAN TRANSMEMBRANE
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Inocyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/839,711
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0270 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 202 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 186804
US-08-839-711-3

Query Match 100.0%; Score 49; DB 3; Length 202;
Best Local Similarity 23.5%; Pred. No. 6.3e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Oy 1 XXXXXCCXXCCXXCX 17
Db 73 DCCCGCGHENGCKRCA 89

Search completed: July 1, 2002, 12:27:59
Job time: 156 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 12:25:58 ; Search time 14.69 Seconds
(without alignments)
111.199 Million cell updates/sec

Title: US-09-493-795A-1
Perfect score: 49
Sequence: 1 XXXXXCCXXCXXCXXCX 17

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

arched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Prd. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 49 | 100.0 | 29 | A55891 | delta-conotoxin Gm |
| 2 | 49 | 100.0 | 55 | S25774 | testis-specific pr |
| 3 | 49 | 100.0 | 64 | A25775 | metallothionein A |
| 4 | 49 | 100.0 | 64 | A33825 | metallothionein A |
| 5 | 49 | 100.0 | 65 | A38739 | metallothionein - |
| 6 | 49 | 100.0 | 66 | S78705 | protein YBR056w-a |
| 7 | 49 | 100.0 | 68 | JU0034 | long neurotoxin ho |
| 8 | 49 | 100.0 | 72 | H42525 | A-ORF-U protein - |
| 9 | 49 | 100.0 | 75 | C71370 | hypothetical prote |
| 10 | 49 | 100.0 | 76 | T09262 | glycine-rich cell |
| 11 | 49 | 100.0 | 78 | T50943 | probable ferredoxi |
| 12 | 49 | 100.0 | 85 | D82813 | hypothetical prote |
| 13 | 49 | 100.0 | 96 | T43470 | hypothetical prote |
| 14 | 49 | 100.0 | 117 | A24178 | Whey acidic protei |
| 15 | 49 | 100.0 | 122 | JC5548 | high sulfur protei |
| 16 | 49 | 100.0 | 126 | I46489 | cysteine-rich hair |
| 17 | 49 | 100.0 | 127 | T34270 | hypothetical prote |
| 18 | 49 | 100.0 | 129 | FTHUB | foliitropin beta c |
| 19 | 49 | 100.0 | 129 | 1 FTSHB | foliitropin beta c |
| 20 | 49 | 100.0 | 129 | 1 FTGBB | foliitropin beta c |
| 21 | 49 | 100.0 | 129 | 2 S3550 | foliitropin beta c |
| 22 | 49 | 100.0 | 131 | S50807 | probable membrane |
| 23 | 49 | 100.0 | 137 | T38500 | hypothetical prote |
| 24 | 49 | 100.0 | 147 | S70109 | hypothetical prote |
| 25 | 49 | 100.0 | 151 | S60314 | hair keratin cyste |
| 26 | 49 | 100.0 | 152 | 1 KRSHHC | keratin high-sulfu |
| 27 | 49 | 100.0 | 152 | 2 I47109 | high-sulfur wool m |
| 28 | 49 | 100.0 | 152 | 2 I47111 | high-sulfur wool m |
| 29 | 49 | 100.0 | 152 | 2 I47112 | high-sulfur wool m |

| | | | | | |
|----|----|-------|-----|----------|--------------------|
| 30 | 49 | 100.0 | 152 | 2 I47108 | high-sulfur wool m |
| 31 | 49 | 100.0 | 152 | 2 T18975 | hypothetical prote |
| 32 | 49 | 100.0 | 156 | 1 KRSHHB | keratin high-sulfu |
| 33 | 49 | 100.0 | 158 | 2 S57452 | ferredoxin 2[4Fe-4 |
| 34 | 49 | 100.0 | 162 | 2 I47107 | high-sulfur wool m |
| 35 | 49 | 100.0 | 164 | 2 T24272 | hypothetical prote |
| 36 | 49 | 100.0 | 166 | 2 C69074 | ferredoxin 2[4Fe-4 |
| 37 | 49 | 100.0 | 169 | 1 S18946 | ultra high-sulfur |
| 38 | 49 | 100.0 | 169 | 2 S51910 | cryptogene protein |
| 39 | 49 | 100.0 | 172 | 1 KRSHHA | keratin high-sulfu |
| 40 | 49 | 100.0 | 172 | 2 I47106 | high-sulfur wool m |
| 41 | 49 | 100.0 | 174 | 2 B95940 | hypothetical prote |
| 42 | 49 | 100.0 | 174 | 2 T21945 | hypothetical prote |
| 43 | 49 | 100.0 | 175 | 2 S37649 | high-sulfur kerati |
| 44 | 49 | 100.0 | 177 | 2 S37650 | high-sulfur kerati |
| 45 | 49 | 100.0 | 182 | 1 KRSHHD | keratin high-sulfu |
| 46 | 49 | 100.0 | 182 | 2 I47105 | high-sulfur wool m |
| 47 | 49 | 100.0 | 183 | 2 T23240 | hypothetical prote |
| 48 | 49 | 100.0 | 186 | 2 A45910 | ultra-high-sulfur |
| 49 | 49 | 100.0 | 188 | 2 JC6547 | high sulfur protei |
| 50 | 49 | 100.0 | 188 | 2 T15651 | hypothetical prote |

ALIGNMENTS

RESULT 1
A55891
delta-conotoxin GmVIA - cone shell (Conus gloriamaris)
C:Species: Conus gloriamaris (glory-of-the-sea cone)
C:Date: 19-Mar-1997 #sequence_revision 11-Apr-1997 #text_change 10-Sep-1997
C:Accession: A55891
R:Shon, K.J.; Hasson, A.; Spira, M.E.; Cruz, L.J.; Gray, W.R.; Olivera, B.M.
Biochemistry 33, 11420-11425, 1994
A:Title: Delta-conotoxin GmVIA, a novel peptide from the venom of Conus gloriamaris.
A:Reference number: A55891; MUID:95001845
A:Accession: A55891
A:Molecule type: protein
A:Residues: 1-29 <SHO>
A>Note: the structure was confirmed by chemical synthesis
C:Comment: This toxin, although similar to the omega-conotoxins which act on calcium
C:Superfamily: unassigned conotoxins
C:Keywords: neurotoxin; sodium channel inhibitor; venom
F:4-19,11-24,18-28/Disulfide bonds: #status experimental

Query Match 100.0%; Score 49; DB 2; Length 29;
Best Local Similarity 23.5%; Pred. No. 2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXCXXCXXCX 17
DB 13 PIFQNCRCGNCVLCV 29

RESULT 2
S25774
testis-specific protein Mst84Dc - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 20-Aug-1999
C:Accession: S25774; C56565
R:Kuhn, R.; Kuhn, C.; Boersch, D.; Glaetzer, K.H.; Schaefer, M.
Mech. Dev. 35, 143-151, 1991
A:Title: A cluster of four genes selectively expressed in the male germ line of Drosophila
A:Reference number: A56565; MUID:92102953
A:Accession: S25774
A:Molecule type: DNA
A:Residues: 1-55 <KUH>
A:Cross-references: EMBL:X67703; NID:g11072; PIDN:CAA47939.1; PID:g11075
A>Note: the authors translated the codon TGC for residue 55 as Thr
A:Note: sequence extracted from NCBI backbone (NCBIN:74217, NCBIIP:74222)
C:Genetics:
A:Gene: Mst84Dc

C;Date: 13-Sep-1991 #sequence_revision 13-Sep-1991 #text_change 20-Aug-1999

Dn 16 SEQELCCTRTWCDQMCQ 32

RESULT 8

H42525
A:ORF-U protein - vaccinia virus (strain Copenhagen)
C:Species: vaccinia virus
A:Note: host Homo sapiens (man)
C:Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 08-Apr-1994
C:Accession: H42525

R:Johnson, G.P.

submitted to GenBank, June 1990

A:Reference number: A33172

A:Accession: H42525

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-72 <JOH>

Query Match 100.0%; Score 49; DB 2; Length 72;

Best Local Similarity 23.5%; Pred. No. 3.2e+02;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17

Db 10 FRSLVCCGFSRCCTVCI 26

RESULT 9

C71370
hypothetical protein TP0059 - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
C:Accession: C71370

R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin

rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McD

they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.

Science 281, 375-388, 1998

A>Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.

A:Reference number: A71250; MUID:98332770

A:Accession: C71370

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-75 <COL>

A:Cross-references: GB:AE001191; GB:AE000520; NID:g3322316; PIDN:AAC65063.1; PID:g332232

A:Experimental source: strain Nichols

C:Genetics:

A:Gene: TP0059

Query Match 100.0%; Score 49; DB 2; Length 75;

Best Local Similarity 23.5%; Pred. No. 3.3e+02;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17

Db 6 SARSMCCSFSCAKNCA 22

RESULT 10

T09262
glycine-rich cell wall protein EMB31 - white spruce
C:Species: Picea glauca (white spruce)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999
C:Accession: T09262

R:Dong, J.Z.; Dunstan, D.I.

submitted to the EMBL Data Library, June 1996

A:Description: Gene expression during somatic embryogenesis.

A:Reference number: Z16588

A:Accession: T09262

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-76 <DON>

A:Cross-references: EMBL:L47748; NID:g1350525; PID:g1350526

C:Genetics:

A:Gene: EMB31

Query Match

100.0%; Score 49; DB 2; Length 76;

Best Local Similarity 23.5%; Pred. No. 3.3e+02;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17

Db 47 YFLCCCLMLQECFGCC 63

RESULT 11

T50943

probable ferredoxin Dita [imported] - Pseudomonas abietaniphila

C:Species: Pseudomonas abietaniphila

C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000

C:Accession: T50943

R:Martin, V.J.; Mohn, W.W.

J. Bacteriol. 181, 2675-2682, 1999

A>Title: A novel aromatic-ring-hydroxylating dioxygenase from the diterpenoid-degradi

A:Reference number: Z25281; MUID:99235742

A:Accession: T50943

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-78 <MAR>

A:Cross-references: EMBL:AF119621; PIDN:AAD21062.1

A:Experimental source: strain BKME-9; ATCC700689

C:Genetics:

A:Gene: dita3

Query Match

100.0%; Score 49; DB 2; Length 78;

Best Local Similarity 23.5%; Pred. No. 3.4e+02;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17

Db 12 ADRSRCCGYGLCAAVCP 28

RESULT 12

D82813

hypothetical protein XF0391 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C:Accession: D82813

R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq

Nature 406, 151-157, 2000

A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: D82813

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-85 <SIM>

A:Cross-references: GB:AE003890; GB:AE003849; NID:g9105215; PIDN:AAF83201.1; GSPDB:GN

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer

as-Neco, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins

A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.

F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, C.

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santeilli, R.V.; Sawa

A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv

M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.

A:Reference number: A59328

A;Contents: annotation
C;Genetics:
A;Gene: XF0391

```
Query Match      100.0%; Score 49; DB 2; Length 85;
Best Local Similarity 23.5%; Pred. No. 3.5e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;
```

C:Species: *Rattus norvegicus* (Norway rat)
C:Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 21-Jul-2000
C:Accession: JC6548
R:Mitsui, S.; Ohuchi, A.; Adachi-Yamada, T.; Hotta, M.; Tsuboi, R.; Ogawa, H.
Gene 208, 123-129, 1998
A:Title: Structure and hair follicle-specific expression of genes encoding the rat hi
A:Reference number: JC6547; MUID:98201605
A:Accession: JC6548
A:Molecule type: DNA
A:Residues: 1-122 <MIT>
A:Cross-references: DDBJ:AB003753; NID:g3046870; PIDN:BAA25574.1; PID:g3046872
C:Comment: This protein is a cysteine-rich, keratin associated protein.
C:Genetics:
A:Gene: b2f
C:Superfamily: keratin high-sulfur matrix protein IIIA

U: Biochem: 139, 193-201, 1968
 Title: A camel milk whey protein rich in half-cystine. Primary structure, assessment of the function of the half-cystine residues and the effect of reduction on the structure
 A: Reference number: A24178; MUID:86300719
 A: Accession: A24178
 A: Molecule type: protein
 A: Residues: 1-117 <BEG>
 C: Superfamily: antileukoproteinas; antileukoproteinas repeat homology
 F: 9-54/Domain: antileukoproteinas repeat homology <ALP2>
 F: 64-111/Domain: antileukoproteinas repeat homology <ALP1>

Query Match 100.0%; Score 49; DB 2; Length 117;
Best Local Similarity 23.5%; Pred. No. 4.2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

R;Wilcox, L.
C;Accession: T34270
A;Submitted to the EMBL Data Library, November 1995
A;Description: The sequence of *C. elegans* cosmid F46C8.
A;Reference number: Z21497
A;Accession: T34270
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-127 <WIL>
A;Cross-references: EMBL:U41624; PIDN:AAA83322.1; CESP:F46C8.1
C;Genetics:
A;Gene: CESP:F46C8.1
A;Introns: 19/2: 56/1: 96/3

A:Gene: GDB:FSHB
 A:Cross-references: GDB:119955; OMIM:136530
 A:Map position: 11p13-11p13
 A:Introns: 53/3
 C:Superfamily: pituitary glycoprotein hormone beta chain
 C:Keywords: glycoprotein; heterodimer; hormone; pituitary
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-129/Product: follitropin beta chain #status experimental <MAT>
 F:21-46,35-69,38-100,50-122,84-112,102-105/Disulfide bonds: #status experimental
 F:25,42/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 100.0%; Score 49; DB 1; Length 129;
 Best Local Similarity 23.5%; Pred No. 4, 4e+02;
 Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
 DB 6 FFELFCCWKAICCSCE 22

RESULT 19

FTSHB
 N:Alternate names: follitropin beta chain precursor - sheep
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C:Date: 18-Dec-1981 #sequence_revision 23-Aug-1996 #text_change 18-Jun-1999
 C:Accession: A40410; S05316; A01495
 R:Guzman, K.; Miller, C.D.; Phillips, C.L.; Miller, W.L.
 DNA Cell Biol. 10, 593-601, 1991
 A:Title: The gene encoding ovine follicle-stimulating hormone beta: isolation, charac
 A:Reference number: A40410; MUID:92029621
 A:Accession: A40410
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-129 <GU>
 A:Cross-references: GB:S64745; NID:9238864; PIDN:AAB20317.1; PID:9238865
 R:Mountford, P.S.; Bello, P.A.; Brandon, M.R.; Adams, T.E.
 Nucleic Acids Res. 17, 6391, 1989
 A:Title: Cloning and DNA sequence analysis of the cDNA for the precursor of ovine fol
 A:Reference number: S05316; MUID:89366671
 A:Accession: S05316
 A:Molecule type: mRNA
 A:Residues: 1-129 <MO>
 A:Cross-references: EMBL:X15493; NID:91251; PIDN:CAA33516.1; PID:9683658
 R:Sairam, M.R.; Seidah, N.G.; Chretien, M.
 Biochem. J. 197, 541-552, 1981
 A:Title: Primary structure of the ovine pituitary follitropin beta-subunit.
 A:Reference number: A01495; MUID:82113053
 A:Accession: A01495
 A:Molecule type: protein
 A:Residues: 20-58, 'B', 60-63, 'B', 65-67, 'T', 69-106, 'S', 108-127, 'ERZ' <SAI>
 C:Superfamily: pituitary glycoprotein hormone beta chain
 C:Keywords: glycoprotein; hormone; pituitary
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-129/Product: follitropin beta chain #status predicted <MAT>
 F:21-46,35-69,38-100,50-122,84-112,102-105/Disulfide bonds: #status predicted
 F:25,42/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 49; DB 1; Length 129;
 Best Local Similarity 23.5%; Pred No. 4, 4e+02;
 Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
 DB 6 FCFLFCCWKAICCSCE 22

RESULT 20

FTPGH
 N:Alternate names: follitropin beta chain precursor - pig
 C:Species: Sus scrofa domestica
 C:Date: 18-Dec-1981 #sequence_revision 23-Aug-1996 #text_change 18-Jun-1999
 C:Accession: A40410; S05316; A01495
 R:Guzman, K.; Miller, C.D.; Phillips, C.L.; Miller, W.L.
 DNA Cell Biol. 10, 593-601, 1991
 A:Title: The gene encoding ovine follicle-stimulating hormone beta: isolation, charac
 A:Reference number: A40410; MUID:92029621
 A:Accession: A40410
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-129 <GU>
 A:Cross-references: GB:S64745; NID:9238864; PIDN:AAB20317.1; PID:9238865
 R:Mountford, P.S.; Bello, P.A.; Brandon, M.R.; Adams, T.E.
 Nucleic Acids Res. 17, 6391, 1989
 A:Title: Cloning and DNA sequence analysis of the cDNA for the precursor of ovine fol
 A:Reference number: S05316; MUID:89366671
 A:Accession: S05316
 A:Molecule type: mRNA
 A:Residues: 1-129 <MO>
 A:Cross-references: EMBL:X15493; NID:91251; PIDN:CAA33516.1; PID:9683658
 R:Sairam, M.R.; Seidah, N.G.; Chretien, M.
 Biochem. J. 197, 541-552, 1981
 A:Title: Primary structure of the ovine pituitary follitropin beta-subunit.
 A:Reference number: A01495; MUID:82113053
 A:Accession: A01495
 A:Molecule type: protein
 A:Residues: 20-58, 'B', 60-63, 'B', 65-67, 'T', 69-106, 'S', 108-127, 'ERZ' <SAI>
 C:Superfamily: pituitary glycoprotein hormone beta chain
 C:Keywords: glycoprotein; hormone; pituitary
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-129/Product: follitropin beta chain #status predicted <MAT>
 F:21-46,35-69,38-100,50-122,84-112,102-105/Disulfide bonds: #status predicted
 F:25,42/Binding site: carbohydrate (Asn) (covalent) #status predicted

C:Species: Sus scrofa domestica (domestic pig)
C:Date: 31-May-1979 #sequence_revision 23-Aug-1997 #text_change 16-Jun-2000
C:Accession: A48169; I46582; A01496
R:Hirai, T.; Takikawa, H.; Kato, Y.
J. Mol. Endocrinol. 5, 147-158, 1990
A:Title: The gene for the beta subunit of porcine FSH: absence of consensus oestrogen-re
A:Reference number: A48169; MUID:91063935
A:Accession: A48169
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-129 <HR>
A:Cross-references: GB:D00621; NID:g217687; PIDN:BAA00499.1; PID:g217688
R:Kato, Y.
Mol. Cell. Endocrinol. 55, 107-112, 1988
A:Title: Cloning and DNA sequence analysis of the cDNA for the precursor of porcine foll
A:Reference number: I46582; MUID:88196589
A:Accession: I46582
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 15-129 <KAT>
A:Cross-references: GB:M35676; NID:g164463; PIDN:AAA31039.1; PID:g164464
R:Closset, J.; Maguin-Rogister, G.; Hennen, G.; Strosberg, A.D.
Eur. J. Biochem. 86, 115-120, 1978
A:Title: Porcine follitropin. The amino-acid sequence of the beta subunit.
A:Reference number: A01496; MUID:78190610
A:Accession: A01496
A:Molecule type: protein
A:Residues: 21,'G',23-30,'GVKGLT',37-50,'T',52,'G',53,'B',55-58,'B',60-63,'B',65,'Z',67-
C:Superfamily: pituitary glycoprotein hormone beta chain
C:Keywords: glycoprotein; heterodimer; hormone; pituitary
F:21-46,35-69,38-100,50-122,84-112,102-105/Disulfide bonds: #status predicted
F:25,42/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 49; DB 1; Length 129;
Best Local Similarity 23.5%; Pred. No. 4.4e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXCX 17
Db 6 FCFLFCWKAICNSCE 22

RESULT 21
A23550
follitropin beta chain precursor - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 23-Aug-1987 #sequence_revision 23-Aug-1987 #text_change 16-Jul-1999
C:Accession: A29816; A23550; A24914
R:Kim, K.E.; Gordon, D.F.; Maurer, R.A.
Proc. Natl. Acad. Sci. U.S.A. 83, 6618-6621, 1986
A:Title: Nucleotide sequence of the bovine gene for follicle-stimulating hormone beta-su
A:Reference number: A29816; MUID:88283341
A:Accession: A29816
A:Molecule type: DNA
A:Residues: 1-129 <KIM>
A:Cross-references: GB:M83753; GB:M20185; NID:g163063; PIDN:AAA30528.1; PID:g163064
R:Esch, F.S.; Mason, A.J.; Cooksey, K.; Mercado, M.; Shimasaki, S.
Proc. Natl. Acad. Sci. U.S.A. 83, 6618-6621, 1986
A:Title: Cloning and DNA sequence analysis of the cDNA for the precursor of the beta cha
A:Reference number: A23550; MUID:86313629
A:Accession: A23550
A:Molecule type: mRNA
A:Residues: 1-129 <ESC>
A:Cross-references: GB:M13383; NID:g163059; PIDN:AAA30526.1; PID:g163060
R:Maurer, R.A.; Beck, A.
DNA 5, 363-369, 1986
A:Title: Isolation and nucleotide sequence analysis of a cloned cDNA encoding the beta-s
A:Reference number: A24914; MUID:87053172
A:Accession: A24914
A:Molecule type: mRNA
A:Residues: 1-129 <MAU>
A:Cross-references: GB:M14853; NID:g163061; PIDN:AAA30527.1; PID:g163062

C:Superfamily: pituitary glycoprotein hormone beta chain
F:21-46,35-69,38-100,50-122,84-112,102-105/Disulfide bonds: #status predicted

Query Match 100.0%; Score 49; DB 2; Length 129;
Best Local Similarity 23.5%; Pred. No. 4.4e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXCX 17
Db 6 FCFLFCWKAICNSCE 22

RESULT 22
S50807
probable membrane protein YJL064w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein HRC131; hypothetical protein J1120
C:Species: Saccharomyces cerevisiae
C:Date: 13-Jan-1995 #sequence_revision 08-Sep-1995 #text_change 21-Jul-2000
C:Accession: S50807; S47126; S56838
R:Vandenbol, M.; Durand, P.; Dion, C.; Portetelle, D.; Hilger, F.
Yeast 11, 57-60, 1995
A:Title: Sequence of a 17.1 kb DNA fragment from chromosome X of Saccharomyces cerevi
A:Reference number: S50798; MUID:95282514
A:Accession: S50807
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-131 <VAN>
A:Cross-references: EMBL:234288; NID:g498992; PIDN:CAA84058.1; PID:g499002
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1994
R:Vandenbol, M.; Durand, P.; Dion, C.; Portetelle, D.; Hilger, F.
submitted to the EMBL Data Library, June 1994
A:Description: Sequence analysis of a 17.1 kb DNA fragment from chromosome X of Sacch
A:Reference number: S47117
A:Accession: S47126
A:Molecule type: DNA
A:Residues: 1-131 <VAV>
A:Cross-references: EMBL:234288; NID:g498992; PID:g499002
R:Vandenbol, M.; Durand, P.; Portetelle, D.; Hilger, F.
submitted to the Protein Sequence Database, September 1995
A:Reference number: S56835
A:Accession: S56838
A:Molecule type: DNA
A:Residues: 1-131 <POH>
A:Cross-references: EMBL:249340; NID:g1008212; PID:g1008214; MIPS:YJL064w
C:Genetics:
A:Map position: 10L
C:Superfamily: Saccharomyces probable membrane protein YJL064w
C:Keywords: transmembrane protein

Query Match 100.0%; Score 49; DB 2; Length 131;
Best Local Similarity 23.5%; Pred. No. 4.5e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXCX 17
Db 55 GTCCGCCCCCLRDSCV 71

RESULT 23
T38500
hypothetical protein SPAC29B12.13 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T38500
R:Gentiles, S.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21797
A:Accession: T38500
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-137 <GEN>

Anim. Genet. 25, 407-415, 1994

A:Reference number: I47105; MUID:95209146
A:Accession: I47111
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-152 <ROG>
A:Cross-references: GB:L33891; NID:g499877; PIDN:AAB01450.1; PID:g940361
A:Accession: I47110
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-152 <ROG>
A:Cross-references: GB:L33890; NID:g499876; PIDN:AAB01449.1; PID:g940360
C:Superfamily: keratin high-sulfur matrix protein IIIA

Query Match 100.0%; Score 49; DB 2; Length 152;
Best Local Similarity 23.5%; Pred. No. 4.9e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 112 YAAASCCRPSCGQSCC 128

RESULT 29
I47112
high-sulfur wool matrix protein B2C - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 15-Sep-2000
C:Accession: I47112
R:Rogers, G.R.; Hickford, J.G.; Bickerstaffe, R.
Anim. Genet. 25, 407-415, 1994
A:Title: Polymorphism in two genes for B2 high sulfur proteins of wool.
A:Reference number: I47105; MUID:95209146
A:Accession: I47112
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-152 <ROG>
A:Cross-references: GB:L33892; NID:g499878; PIDN:AAB01451.1; PID:g940362
C:Superfamily: keratin high-sulfur matrix protein IIIA

Query Match 100.0%; Score 49; DB 2; Length 152;
Best Local Similarity 23.5%; Pred. No. 4.9e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 112 YAAASCCRPSCGQSCC 128

SULT 30
I47108
high-sulfur wool matrix protein B2C - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 15-Sep-2000
C:Accession: I47108
R:Rogers, G.R.; Hickford, J.G.; Bickerstaffe, R.
Anim. Genet. 25, 407-415, 1994
A:Title: Polymorphism in two genes for B2 high sulfur proteins of wool.
A:Reference number: I47105; MUID:95209146
A:Accession: I47108
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-152 <ROG>
A:Cross-references: GB:L33888; NID:g499874; PIDN:AAB01447.1; PID:g940358
C:Superfamily: keratin high-sulfur matrix protein IIIA

Query Match 100.0%; Score 49; DB 2; Length 152;
Best Local Similarity 23.5%; Pred. No. 4.9e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17

DB 112 YAAASCCRPSCGQSCC 128

RESULT 31
T18975
hypothetical protein C06A1.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18975
R:McMurray, A.
submitted to the EMBL Data Library, June 1995
A:Reference number: Z19054
A:Accession: T18975
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-152 <WIL>
A:Cross-references: EMBL:Z49886; PIDN:CAA90055.1; GSPDB:GN00020; CESP:C06A1.6
A:Experimental source: clone C06A1
C:Genetics:
A:Gene: CESP:C06A1.6
A:Map position: 2
A:Introns: 22/3

Query Match 100.0%; Score 49; DB 2; Length 152;
Best Local Similarity 23.5%; Pred. No. 4.9e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 67 GCGCCCRPRCCCCR 83

RESULT 32
KRSHHB
keratin high-sulfur matrix protein B2B - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 24-Apr-1984 #sequence_revision 24-Apr-1984 #text_change 16-Aug-1996
C:Accession: A02838
R:Elleman, T.C.; Doppeide, T.A.
J. Biol. Chem. 247, 3900-3909, 1972
A:Title: The sequence of SCMK-B2B, a high sulfur protein from wool keratin.
A:Reference number: A02838; MUID:72206034
A:Accession: A02838
A:Molecule type: protein
A:Residues: 1-156 <ELL>
A:Experimental source: Lincoln wool
A:Note: a minor component has 77-Gly
C:Comment: Most of the sequence shows a repeating pattern of 10 residues.
C:Superfamily: keratin high-sulfur matrix protein IIIA
C:Keywords: acetylated amino end; duplication; hair
F:1/Modified site: acetylated amino end (Ala) #status experimental

Query Match 100.0%; Score 49; DB 1; Length 156;
Best Local Similarity 23.5%; Pred. No. 4.9e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 121 YAAASCCRPSCGQSCC 137

RESULT 33
S57452
ferredoxin 2[4Fe-4S] fwdE - Methanobacterium thermoautotrophicum (strain Marburg)
C:Species: Methanobacterium thermoautotrophicum
C:Date: 10-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 20-Jun-2000
C:Accession: S63541; S57452
R:Hochheimer, A.; Schmitz, R.A.; Thauer, R.K.; Hedderich, R.
Eur. J. Biochem. 234, 910-920, 1995
A:Title: The tungsten formylmethanofuran dehydrogenase from Methanobacterium thermoau

A:Reference number: S63519; MUID:96163477

A:Accession: S63541

A:Molecule type: DNA

A:Residues: 1-158 <HOC>

A:Cross-references: EMBL:X87970; NID:gl890205; PIDN:CAA61209.1; PID:g1890206

A:Experimental source: strain Marburg, DSM 2133

C:Genetics:

A:Gene: fwdE

C:Superfamily: Methanobacterium ferredoxin 2[4Fe-4S] fwdE; ferredoxin 2[4Fe-4S] homolog

C:Keywords: 4Fe-4S; iron-sulfur protein; metalloprotein

F:92-147/Domain: ferredoxin 2[4Fe-4S] homolog <PER>

F:100,103,106,139/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

F:110,129,132,135/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 100.0%; Score 49; DB 2; Length 158;

Best Local Similarity 23.5%; Pred. No. 5e+02;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

1 XXXXXCCXXXCCXXXCX 17

DB 124 VDSEWCCGCLICVDMCP 140

100.0%; Score 49; DB 2; Length 158;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

RESULT 34

I47107

high-sulfur wool matrix protein B2A - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C>Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 15-Sep-2000

C:Accession: I47107

R:Rogers, G.R.; Hickford, J.G.; Bickerstaffe, R.

Anim. Genet. 25, 407-415, 1994

A:Title: Polymorphism in two genes for B2 high sulfur proteins of wool.

A:Reference number: I47105; MUID:95209146

A:Accession: I47107

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-162 <ROG>

A:Cross-references: GB:L33887; NID:g499873; PIDN:AAB01446.1; PID:g940357

C:Superfamily: Keratin high-sulfur matrix protein IIIA

Query Match 100.0%; Score 49; DB 2; Length 162;

Best Local Similarity 23.5%; Pred. No. 5e+02;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

1 XXXXXCCXXXCCXXXCX 17

DB 122 YQAASCCRPSCGQSCC 138

100.0%; Score 49; DB 2; Length 162;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

RESULT 35

T24272

hypothetical protein T01B7.8 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T24272

R:Sim, M.

submitted to the EMBL Data Library, October 1995

A:Reference number: 219867

A:Accession: T24272

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-164 <WIL>

A:Cross-references: EMBL:Z66499; PIDN:CAA91301.1; GSPDB:GN00020; CESP:T01B7.8

A:Experimental source: clone T01B7

C:Genetics:

A:Gene: CESP:T01B7.8

A:Map position: 2

A:Introns: 20/3; 90/2

Query Match 100.0%; Score 49; DB 2; Length 164;

Best Local Similarity 23.5%; Pred. No. 5.1e+02;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17

DB 80 GCGGCCCRPRCCGCC 96

100.0%; Score 49; DB 2; Length 164;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

RESULT 36

C69074

ferredoxin 2[4Fe-4S] fwdE - Methanobacterium thermoautotrophicum (strain Delta H)

N:Alternate names: tungsten formylmethanofuran dehydrogenase subunit H [mismomer]

C:Species: Methanobacterium thermoautotrophicum

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 04-Feb-2000

C:Accession: C69074

R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.

Oiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani,

Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu

A:Reference number: A69000; MUID:98037514

A:Accession: C69074

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-166 <WTH>

A:Cross-references: GB:AE000916; GB:AE000666; NID:g2622674; PIDN:AAB86027.1; PID:g262

A:Experimental source: strain Delta H

A>Note: Met-8 may be in a better context to be the initiation codon

C:Genetics:

A:Gene: MTH1553

A:Start codon: GTG

C:Superfamily: Methanobacterium ferredoxin 2[4Fe-4S] fwdE; ferredoxin 2[4Fe-4S] homol

C:Keywords: 4Fe-4S; iron-sulfur protein; metalloprotein

F:102-155/Domain: ferredoxin 2[4Fe-4S] homolog <PER>

F:108,111,114,147/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

F:118,137,140,143/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 100.0%; Score 49; DB 2; Length 165;

Best Local Similarity 23.5%; Pred. No. 5.1e+02;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17

DB 132 IESEWCCGCLICVDMCP 148

100.0%; Score 49; DB 2; Length 165;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

RESULT 37

S18946

ultra high-sulfur keratin 1 - human

N:Alternate names: UHS keratin; ultra high-sulfur matrix protein

C:Species: Homo sapiens (man)

C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C:Accession: S18946; B36686

R:Drabant, B.; Doenecke, D.

submitted to the EMBL Data Library, December 1991

A:Description: Nucleotide sequence of a Human high-sulphur keratin cDNA.

A:Reference number: S18946

A:Accession: S18946

A:Molecule type: mRNA

A:Residues: 1-169 <DRA>

A:Cross-references: EMBL:X63755; NID:g32471; PIDN:CAA45283.1; PID:g32472

R:MacKinnon, P.J.; Powell, B.C.; Rogers, G.E.

J. Cell Biol. 111, 2587-2600, 1990

A:Title: Structure and expression of genes for a class of cysteine-rich proteins of t

A:Reference number: A36686; MUID:91115951

A:Accession: B36686

A:Molecule type: DNA

A:Residues: 1-39, 'y', 41-169 <MAC>

A:Cross-references: GB:X55293; NID:g34078; PIDN:CAA39005.1; PID:g34079

C:Genetics:

A:Gene: GDB:KRNI

A:Cross-references: GDB:125257; OMIM:148021

Biochem. J. 130, 833-845, 1972
 A>Title: The amino acid sequence of protein SCMK-B2A from the high-sulphur fraction of w
 A:Reference number: A02837; MUID:73224964
 A:Accession: A02837

AA:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, A.A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
AA:Reference number: A96039; PMID:21368234; PMID:11474104

A:Contents: annotation
C:Genetics:
A:Gene: SMB21282
A:Genome: plasmid

Query Match 100.0%; Score 49; DB 2; Length 174;
Best Local Similarity 23.5%; Pred. No. 5.2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 39 ISTRGCLRRRCRLPCP 55

RESULT 42

T21945
Hypothetical protein F38B2.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T21945

R:Kershaw, J.

submitted to the EMBL Data Library, July 1995

A:Reference number: Z19492

A:Accession: T21945

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-174 <WIL>

A:Cross-references: EMBL:Z50045; PIDN:CAA90362.1; GSPDB:GN00028; CBSP:F38B2.2

A:Experimental source: clone F38B2

C:Genetics:

A:Gene: CBSP:F38B2.2

A:Map position: X

A:Introns: 41/1; 81/3; 125/3

Query Match 100.0%; Score 49; DB 2; Length 174;
Best Local Similarity 23.5%; Pred. No. 5.2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 80 LQDSFCCSEGPCLTRCG 96

RESULT 43

S37649
high-sulfur keratin - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Aug-1999
C:Accession: S37649
R:Zhumabaeva, B.D.; Gening, L.V.; Gazaryan, K.G.
Mol. Biol. 26, 550-555, 1992
A:Title: Cloning and structural characterization of human hair sulfur-rich keratin genes
A:Reference number: S37649
A:Accession: S37649
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-175 <ZHU>
A:Cross-references: EMBL:X63338; MID:g311881; PIDN:CAA44938.1; PID:g311882
C:Superfamily: keratin high-sulfur matrix protein IIIA

Query Match 100.0%; Score 49; DB 2; Length 175;
Best Local Similarity 23.5%; Pred. No. 5.2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 19 TCGSSCCQPCCTSCC 35

RESULT 44

S37650
high-sulfur keratin - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Aug-1999
C:Accession: S37650
R:Zhumabaeva, B.D.; Gening, L.V.; Gazaryan, K.G.
Mol. Biol. 26, 550-555, 1992
A:Title: Cloning and structural characterization of human hair sulfur-rich keratin ge
A:Reference number: S37649
A:Accession: S37650
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-177 <ZHU>
A:Cross-references: EMBL:X63337; MID:g311879; PIDN:CAA44937.1; PID:g311880
C:Superfamily: keratin high-sulfur matrix protein IIIA

Query Match 100.0%; Score 49; DB 2; Length 177;
Best Local Similarity 23.5%; Pred. No. 5.3e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 24 CCQPCCTSCCQPCRC 40

RESULT 45

KRSHHD
keratin high-sulfur matrix protein B2D - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 22-Jun-1999
C:Accession: S07911
R:Powell, B.C.; Sleight, M.J.; Ward, K.A.; Rogers, G.E.
Nucleic Acids Res. 11, 5327-5346, 1983
A:Title: Mammalian keratin gene families: organisation of genes coding for the B2 hig
A:Reference number: S07349; MUID:83299218
A:Accession: S07911
A:Molecule type: DNA
A:Residues: 1-182 <POW>
A:Cross-references: EMBL:X01610; MID:g1295; PIDN:CAA25759.1; PID:g1298
C:Superfamily: keratin high-sulfur matrix protein IIIA
C:Keywords: duplication; hair
F:2-182/Product: keratin high-sulfur matrix protein B2D #status predicted <MAT>
F:27-36,37-46,47-56,57-66,67-76,77-86/Region: duplication

Query Match 100.0%; Score 49; DB 1; Length 182;
Best Local Similarity 23.5%; Pred. No. 5.4e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 142 YAOACCRPSYCGQSCC 158

RESULT 46

I47105
high-sulfur wool matrix protein B2A - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 15-Sep-2000
C:Accession: I47105
R:Rogers, G.R.; Hickford, J.G.; Bickertstaffe, R.
Anim. Genet. 25, 407-415, 1994
A:Title: Polymorphism in two genes for B2 high sulfur proteins of wool.
A:Reference number: I47105; MUID:95209146
A:Accession: I47105
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-182 <ROG>
A:Cross-references: GB:I33885; MID:g499871; PIDN:AA01444.1; PID:g940355
C:Superfamily: keratin high-sulfur matrix protein IIIA

Query Match 100.0%; Score 49; DB 2; Length 182;
Best Local Similarity 23.5%; Pred. No. 5.4e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCCXXXCCXXXCX 17
Db 142 YAAQACCRPSYQGSCC 158

RESULT 47
T23240
hypothetical protein K02D3.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T23240
R:Bardill, S.
submitted to the EMBL Data Library, March 1996
A:Reference number: 219714
A:Accession: T23240
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-183 <WIL>
A:Cross-references: EMBL:Z70209; PTDN:CAA94146.1; GSPDB:GN00028; CESP:K02D3.1
A:Experimental source: clone K02D3
C:Genetics:
A:Gene: CESP:K02D3.1
A:Map position: X
A:Introns: 47/1; 87/3; 135/3

Query Match 100.0%; Score 49; DB 2; Length 183;
Best Local Similarity 23.5%; Pred. No. 5.4e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCCXXXCCXXXCX 17
Db 86 FOEQFCCASECLARCN 102

RESULT 48
A45910
ultra-high-sulfur keratin - mouse
C:Species: Mus musculus (house mouse)
C:Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 24-Sep-1999
C:Accession: A45910
R:McNab, A.R.; Wood, L.; Theriault, N.; Gierman, T.; Vogeli, G.
J. Invest. Dermatol. 92, 263-266, 1989
A:Title: An ultra-high sulfur keratin gene is expressed specifically during hair growth.
A:Reference number: A45910; MUID:89140394
A:Accession: A45910
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-186 <MCN>
A:Cross-references: GB:M27685; NID:g341749; PTDN:AAA81560.1; PID:g1066818
C:Superfamily: ultra-high-sulfur keratin

Query Match 100.0%; Score 49; DB 2; Length 186;
Best Local Similarity 23.5%; Pred. No. 5.4e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCCXXXCCXXXCX 17
Db 4 SCCSPCCQPTCCRTCC 20

RESULT 49
JC6547
high sulfur protein B2E - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 21-Jul-2000
C:Accession: JC6547
R:Mitsui, S.; Ohuchi, A.; Adachi-Yamada, T.; Hotta, M.; Tsuboi, R.; Ogawa, H.

Gene 208, 123-129, 1998
A:Title: Structure and hair follicle-specific expression of genes encoding the rat hi
A:Reference number: JC6547; MUID:98201605
A:Accession: JC6547
A:Molecule type: DNA
A:Residues: 1-188 <MIT>
A:Cross-references: DDBJ:AB003753; NID:g3046870; PTDN:BAA25573.1; PID:g3046871
C:Comment: This protein is a cysteine-rich, keratin associated protein.
C:Genetics:
A:Gene: b2E
C:Superfamily: keratin high-sulfur matrix protein IIIA
C:Keywords: hair

Query Match 100.0%; Score 49; DB 2; Length 188;
Best Local Similarity 23.5%; Pred. No. 5.5e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCCXXXCCXXXCX 17
Db 69 CSQSSCCQPSCCQTSCC 85

RESULT 50
T15651
hypothetical protein C27A2.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15651
R:Nhan, M.
submitted to the EMBL Data Library, May 1996
A:Description: The sequence of C. elegans cosmid C27A2.
A:Reference number: Z18382
A:Accession: T15651
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-188 <NHA>
A:Cross-references: EMBL:U58760; NID:g1330384; PID:g1330389; PTDN:AAB00710.1; GSPDB:G
A:Experimental source: strain Bristol N2; clone C27A2
C:Genetics:
A:Gene: CESP:C27A2.5
A:Map position: 2
A:Introns: 19/3; 91/2

Query Match 100.0%; Score 49; DB 2; Length 188;
Best Local Similarity 23.5%; Pred. No. 5.5e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCCXXXCCXXXCX 17
Db 81 GCGGCCCRPKCCCCCR 97

Search completed: July 1, 2002, 12:28:21
Job time: 143 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 12:27:08 ; Search time 10.31 seconds
(without alignments)
63.844 Million cell updates/sec

Title: US-09-493-795A-1
Perfect score: 49
Sequence: 1 XXXXXCXXXXXXCXXCX 17

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|--------------|---------------------|
| 1 | 49 | 100.0 | 30 | 1 TX2_HETVE | P58426 heteropoda |
| 2 | 49 | 100.0 | 31 | 1 TX3_HETVE | P58427 heteropoda |
| 3 | 49 | 100.0 | 55 | 1 M84C_DROME | Q01644 drosophila |
| 4 | 49 | 100.0 | 64 | 1 MTA_STRPU | P04734 strongyloce |
| 5 | 49 | 100.0 | 64 | 1 MT_STENE | P55953 sterechius |
| 6 | 49 | 100.0 | 65 | 1 MTB_STRPU | Q27287 strongyloce |
| 7 | 49 | 100.0 | 68 | 1 NXLH_PSEAU | P14612 pseudochis |
| 8 | 49 | 100.0 | 70 | 1 CX2X_CONBE | Q9u323 conus betul |
| 9 | 49 | 100.0 | 72 | 1 YVAU_VAGCC | P20530 vaccinia vi |
| 10 | 49 | 100.0 | 75 | 1 Y059_TREPA | O83098 treponema p |
| 11 | 49 | 100.0 | 84 | 1 HEPG_HUMAN | P81172 homo sapien |
| 12 | 49 | 100.0 | 96 | 1 YZX6_HUMAN | Q9uf47 homo sapien |
| 13 | 49 | 100.0 | 117 | 1 WAP_CAMDR | P09837 camelus dro |
| 14 | 49 | 100.0 | 129 | 1 FSHB_BOVIN | P04837 bos taurus |
| 15 | 49 | 100.0 | 129 | 1 FSHB_CAVPO | Q9jk69 cavia porce |
| 16 | 49 | 100.0 | 129 | 1 FSHB_HORSE | P01226 equus cabal |
| 17 | 49 | 100.0 | 129 | 1 FSHB_HUMAN | P01225 homo sapien |
| 18 | 49 | 100.0 | 129 | 1 FSHB_PIG | P01228 sus scrofa |
| 19 | 49 | 100.0 | 129 | 1 FSHB_SHEEP | P01227 ovis aries |
| 20 | 49 | 100.0 | 131 | 1 YJG4_YEAST | P47038 saccharomyc |
| 21 | 49 | 100.0 | 132 | 1 WAP_PIG | O45655 sus scrofa |
| 22 | 49 | 100.0 | 151 | 1 KR2C_SHEEP | P02440 ovis aries |
| 23 | 49 | 100.0 | 156 | 1 KR2B_SHEEP | P02439 ovis aries |
| 24 | 49 | 100.0 | 169 | 1 KRUA_HUMAN | P26371 homo sapien |
| 25 | 49 | 100.0 | 171 | 1 KR2A_SHEEP | P02438 ovis aries |
| 26 | 49 | 100.0 | 181 | 1 KR2D_SHEEP | P08131 ovis aries |
| 27 | 49 | 100.0 | 194 | 1 KRUB_HUMAN | O75690 homo sapien |
| 28 | 49 | 100.0 | 195 | 1 CSP_TORCA | P56101 torpedo cal |
| 29 | 49 | 100.0 | 197 | 1 CSP_XENLA | O42196 xenopus lae |
| 30 | 49 | 100.0 | 197 | 1 T4S5_HUMAN | O14894 homo sapien |
| 31 | 49 | 100.0 | 198 | 1 DJC5_BOVIN | Q29455 bos taurus |
| 32 | 49 | 100.0 | 198 | 1 DJC5_HUMAN | Q9h324 homo sapien |
| 33 | 49 | 100.0 | 198 | 1 DJC5_MOUSE | P54101 mus musculus |

| | | | | | |
|----|----|-------|-----|--------------|---------------------|
| 34 | 49 | 100.0 | 202 | 1 T4S1_HUMAN | P30408 homo sapien |
| 35 | 49 | 100.0 | 210 | 1 RGSJ_CHICK | Q9pwa0 gallus gall |
| 36 | 49 | 100.0 | 210 | 1 RGSJ_HUMAN | Q9ugc6 homo sapien |
| 37 | 49 | 100.0 | 210 | 1 RGSJ_MOUSE | Q9azb0 mus musculus |
| 38 | 49 | 100.0 | 216 | 1 RGSJ_MOUSE | Q9cx84 mus musculus |
| 39 | 49 | 100.0 | 216 | 1 RGSJ_RAT | O70521 rattus norv |
| 40 | 49 | 100.0 | 217 | 1 RGSJ_HUMAN | P49795 homo sapien |
| 41 | 49 | 100.0 | 218 | 1 RGSJ_CHICK | Q9pwa1 gallus gall |
| 42 | 49 | 100.0 | 239 | 1 RGSJ_MOUSE | Q9azb1 mus musculus |
| 43 | 49 | 100.0 | 245 | 1 IE0_NPVOP | O10369 oryza pseu |
| 44 | 49 | 100.0 | 249 | 1 CSP_DROME | Q03751 drosophila |
| 45 | 49 | 100.0 | 261 | 1 IE0_NPVAC | P41710 autographa |
| 46 | 49 | 100.0 | 267 | 1 APHC_HUMAN | Q9nun7 homo sapien |
| 47 | 49 | 100.0 | 267 | 1 APHC_MOUSE | O94099 mus musculus |
| 48 | 49 | 100.0 | 325 | 1 Y856_TREPA | O83828 treponema p |
| 49 | 49 | 100.0 | 374 | 1 RGSJ_BOVIN | P79348 bos taurus |
| 50 | 49 | 100.0 | 388 | 1 RGSJ_HUMAN | O76081 homo sapien |

ALIGNMENTS

RESULT 1

TX2_HETVE
ID TX2_HETVE STANDARD: PRT: 30 AA.

AC P58426;

DT 01-MAR-2002 (Rel. 41, Created)

DT 01-MAR-2002 (Rel. 41, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Heteropodatoxin 2 (HPTX2).

OS Heteropoda venatoria (Giant crab spider).

OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;

OC Araneomorphae; Entelegynae; Dionycha; Sparassidae; Heteropoda.

OX NCBI_TaxID=152925;

RN [1]

RP SEQUENCE, CHARACTERIZATION, AND MASS SPECTROMETRY.

RC TISSUE=Venom;

RX MEDLINE=97211638; PubMed=9058605;

RA Saquinetti M.C., Johnson J.H., Hammerland L.G., Kelbaugh P.R.,

RA Volkman R.A., Saccomano N.A., Mueller A.L.;

RT "Heteropodatoxins: peptides isolated from spider venom that block

RV K4.2 potassium channels.";

RL Mol. Pharmacol. 51:491-498(1997).

RN [2]

RP STRUCTURE BY NMR.

RC TISSUE=Venom;

RX MEDLINE=21025439; PubMed=11152117;

RA Bernard C., Legros C., Ferrat G., Bischoff U., Marquardt A., Pongs O.,

RA Darbon H.;

RT "Solution structure of hPTX2, a toxin from Heteropoda venatoria

RV spider that blocks Kv4.2 potassium channel.";

RL Protein Sci. 9:2059-2067(2000).

CC -!- FUNCTION: Inhibitor of voltage-gated potassium channels. Blocks

CV potassium currents by binding to Kv4.2 potassium channels.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Produced by the venomous gland.

CC -!- MASS SPECTROMETRY: MW=3412.72; METHOD=Electrospray.

CC -!- SIMILARITY: BELONGS TO THE SPIDER POTASSIUM CHANNEL INHIBITORY

CV TOXIN FAMILY.

DR PDB; LEMX; 24-JAN-01.

KW Toxin; Neurotoxin; Potassium channel inhibitor; Amidation;

KW 3D-structure.

FT DISULFID 3 17

FT DISULFID 10 22

FT DISULFID 16 26

FT MOD_RES 30 30

SQ SEQUENCE 30 AA; 3420 MW; F2A1DC16B695CFCF CRC64;

AMIDATION.

Query Match 100.0%; Score 49; DB 1; Length 30;

Best Local Similarity 23.5%; Pred. No. 47;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
 Db 11 DTNADCEGVVCLWCR 27

RESULT 2

TX3_HETVE STANDARD; PRT; 31 AA.
 AC P58427;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Heteropodatoxin 3 (HPTX3).
 OS Heteropoda venatoria (Giant crab spider).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
 OC Araneomorphae; Entelegynae; Dionycha; Sparassidae; Heteropoda.
 ON NCBI_TaxID=152925;
 [1]
 RP SEQUENCE, CHARACTERIZATION, AND MASS SPECTROMETRY.
 RC TISSUE=Venom;
 RC MEDLINE=97211638; PubMed=9058605;
 RA Sanguinetti M.C., Johnson J.H., Hammerland L.G., Kelbaugh P.R.,
 RA Volkman R.A., Saccomano N.A., Mueller A.B.;
 RT "Heteropodatoxins: peptides isolated from spider venom that block
 RT Kv4.2 potassium channels.";
 RL Mol. Pharmacol. 51:491-498(1997).
 CC -1- FUNCTION: Inhibitor of voltage-gated potassium channels. Blocks
 CC potassium currents by binding to Kv4.2 potassium channels.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Produced by the venomous gland.
 CC -1- MASS SPECTROMETRY: MW=3599.38; METHOD=Electrospray.
 CC -1- SIMILARITY: BELONGS TO THE SPIDER POTASSIUM CHANNEL INHIBITORY
 CC TOXIN FAMILY.
 KW Toxin; Neurotoxin; Potassium channel inhibitor; Amidation.
 FT DISULFID 2 16 BY SIMILARITY.
 FT DISULFID 9 21 BY SIMILARITY.
 FT DISULFID 15 25 BY SIMILARITY.
 FT MOD_RES 31 31 AMIDATION.
 SQ SEQUENCE 31 AA; 3606 MW; BE2558D2E1053095 CRC64;

Query Match 100.0%; Score 49; DB 1; Length 31;
 Best Local Similarity 23.5%; Pred. No. 48;
 Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
 Db 10 STHADCEGFICKLWCR 26

RESULT 3

M84C_DROME STANDARD; PRT; 55 AA.
 AC Q01644; Q9VIA0;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Male specific sperm protein Mst84Dc.
 GN Mst84Dc OR Cg17945.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 ON NCBI_TaxID=7227;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OREGON-R;
 RC MEDLINE=92102953; PubMed=1684716;
 RA Kuhn R., Kuhn C., Boersch D., Glaetzer K.H., Schaefer U.,
 RA Schaefer M.;
 RT "A cluster of four genes selectively expressed in the male germ line
 RT of Drosophila melanogaster.";
 RL Mech. Dev. 35:143-151(1991).

[2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Gaig N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapieton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -1- TISSUE SPECIFICITY: TESTIS.
 CC -1- DEVELOPMENTAL STAGE: PRIMARY SPERMATOCYTES.
 CC -1- DOMAIN: THIS PROTEIN IS MOSTLY COMPOSED OF REPETITIVE C-G-P
 CC MOTIFS.
 CC -1- SIMILARITY: BELONGS TO THE MST(3)CGP FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X67703; CAA47939.1; -;
 DR HSSP; P01180; INPO.
 DR FlyBase; Fgn0004174; Mst84Dc.
 DR FlyBase; Fgn0004174; Mst84Dc.
 DR FlyBase; Fgn0004174; Mst84Dc.
 KW Spermatogenesis; Repeat; Multigene family.
 SQ SEQUENCE 55 AA; 5225 MW; 95A12F3AEC88BD6C CRC64;

Query Match 100.0%; Score 49; DB 1; Length 55;
 Best Local Similarity 23.5%; Pred. No. 70;
 Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
 Db 4 GPGSCCGYCCGPGCG 20


```

RESULT 4
MTA_STRPU
ID MTA_STRPU STANDARD; PRT; 64 AA.
AC P04734;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-NOV-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Metallothionein-A (MTA).
GN MTA.
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90066495; PubMed=2586524;
RA Harlow P., Watkins E., Thornton R.D., Nemer M.;
RT "Structure of an ectodermally expressed sea urchin metallothionein
RT gene and characterization of its metal-responsive region.";
RL Mol. Cell. Biol. 9:5445-5455(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85270452; PubMed=3860837;
RA Nemer M., Wilkinson D.G., Travaglini E.C., Sternberg E.J., Butt T.R.;
RT "Sea urchin metallothionein sequence: key to an evolutionary
RT diversity";
RL Proc. Natl. Acad. Sci. U.S.A. 82:4992-4994(1985).
RN [3]
RP STRUCTURE BY NMR.
RX MEDLINE=99370086; PubMed=10438629;
RA Riek R., Precheur B., Wang Y., Mackay E.A., Wider G., Guntert P.,
RA Liu A., Kaegi J.H.R., Wuthrich K.;
RT "NMR structure of the sea urchin (Strongylocentrotus purpuratus)
RT metallothionein MTA.";
RL J. Mol. Biol. 291:417-428(1999).
CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
CC RESIDUES THAT BIND VARIOUS HEAVY METALS.
CC -!- INDUCTION: BY HEAVY METALS.
CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 4.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M30606; AAA30067.1; -
DR EMBL; K02464; AAA30061.1; -
DR PIR; A33825; A33825.
DR PDB; 1QJL; 31-AUG-99.
DR PDB; 1QJL; 31-AUG-99.
DR InterPro; IPR001396; Metallothion_Echnd.
DR PRINTS; PR00873; MTECHNOIDEA.
KW Metal-binding; Metal-thiolate cluster; Chelation.
FT CONFLICT 10 20 K -> T (IN REF. 2).
FT CONFLICT 24 24 K -> V (IN REF. 2).
SQ SEQUENCE 64 AA; 6444 MW; 56F0AA7A4991E3E6 CRC64;

Query Match 100.0%; Score 49; DB 1; Length 64;
Best Local Similarity 23.5%; Pred. No. 78;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCXXCCXXCCXXCX 17
Db 23 CKTGECCKDGTCCGIC 39

RESULT 5
MTA_STRPU
ID MTA_STRPU STANDARD; PRT; 64 AA.
AC P04734;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-NOV-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Metallothionein-A (MTA).
GN MTA.
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90066495; PubMed=2586524;
RA Harlow P., Watkins E., Thornton R.D., Nemer M.;
RT "Structure of an ectodermally expressed sea urchin metallothionein
RT gene and characterization of its metal-responsive region.";
RL Mol. Cell. Biol. 9:5445-5455(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85270452; PubMed=3860837;
RA Nemer M., Wilkinson D.G., Travaglini E.C., Sternberg E.J., Butt T.R.;
RT "Sea urchin metallothionein sequence: key to an evolutionary
RT diversity";
RL Proc. Natl. Acad. Sci. U.S.A. 82:4992-4994(1985).
RN [3]
RP STRUCTURE BY NMR.
RX MEDLINE=99370086; PubMed=10438629;
RA Riek R., Precheur B., Wang Y., Mackay E.A., Wider G., Guntert P.,
RA Liu A., Kaegi J.H.R., Wuthrich K.;
RT "NMR structure of the sea urchin (Strongylocentrotus purpuratus)
RT metallothionein MTA.";
RL J. Mol. Biol. 291:417-428(1999).
CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
CC RESIDUES THAT BIND VARIOUS HEAVY METALS.
CC -!- INDUCTION: BY HEAVY METALS.
CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 4.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M30606; AAA30067.1; -
DR EMBL; K02464; AAA30061.1; -
DR PIR; A33825; A33825.
DR PDB; 1QJL; 31-AUG-99.
DR PDB; 1QJL; 31-AUG-99.
DR InterPro; IPR001396; Metallothion_Echnd.
DR PRINTS; PR00873; MTECHNOIDEA.
KW Metal-binding; Metal-thiolate cluster; Chelation; 3D-structure.
FT CONFLICT 10 20 K -> T (IN REF. 2).
FT CONFLICT 24 24 K -> V (IN REF. 2).
SQ SEQUENCE 64 AA; 6444 MW; 56F0AA7A4991E3E6 CRC64;

Query Match 100.0%; Score 49; DB 1; Length 64;
Best Local Similarity 23.5%; Pred. No. 78;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCXXCCXXCCXXCX 17
Db 23 CKTGECCKDGTCCGIC 39

RESULT 5
MT_STENE

```

```

ID MT_STENE STANDARD; PRT; 64 AA.
AC P5953;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Metallothionein (MT).
OS Sterechninus neumayeri (Antarctic sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Echinidae;
OC Sterechninus.
OX NCBI_TaxID=53479;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97319675; PubMed=9176568;
RA Scudiero R., Capasso C., Carginale V., Riggio M., Capasso A.,
RA Ciaramella M., Filosa S., Parisi E.;
RT "PCR amplification and cloning of metallothionein complementary DNAs
RT in temperate and Antarctic sea urchin characterized by a large
RT difference in egg metallothionein content.";
RL Cell. Mol. Life Sci. 53:472-477(1997).
CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
CC RESIDUES THAT BIND VARIOUS HEAVY METALS.
CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 4.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y08621; CAA69912.1; -
DR HSSP; P04734; 1QJL.
DR InterPro; IPR001396; Metallothion_Echnd.
DR PRINTS; PR00873; MTECHNOIDEA.
KW Metal-binding; Metal-thiolate cluster; Chelation.
SQ SEQUENCE 64 AA; 6428 MW; 326558A25CA0BAA CRC64;

Query Match 100.0%; Score 49; DB 1; Length 64;
Best Local Similarity 23.5%; Pred. No. 78;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCXXCCXXCCXXCX 17
Db 23 CKTGECCKDGTCCGIC 39

RESULT 6
MTB_STRPU
ID MTB_STRPU STANDARD; PRT; 65 AA.
AC Q27287;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Metallothionein-B (MTB).
GN MTB1 OR MTB.
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87172743; PubMed=3561398;
RA Wilkinson D.G., Nemer M.;
RT "Metallothionein genes Mta and Mtb expressed under distinct
RT quantitative and tissue-specific regulation in sea urchin embryos.";
RL Mol. Cell. Biol. 7:48-58(1987).
RN [2]
RP SEQUENCE FROM N.A.

```

```

Query Match      100.0%; Score 49; DB 1; Length 68;
Best Local Similarity 23.5%; Pred.NO. 81;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY      1 XXXXXCCXXXCCXXXCX 17
        :|::||::|::|::|::|:
Db       16 SEQLCCTKTWCDQCQ 32

RESULT      8
CX2X_CONBE STANDARD; PRT; 70 AA.
ID CX2X_CONBE STANDARD; PRT; 70 AA.
AC Q9U3Z3;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Conotoxin beta precursor.
OS Conus betulinus (Beech cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus..
OX NCBI_TaxID=89764;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RA Fan C.-X.;

```

```

000000 south China sea."
000001 RL Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
000002 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
000003 -----
000004 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
000005 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
000006 CC at the European Bioinformatics Institute. There are no restrictions on its
000007 CC use by non-profit institutions as long as its content is in no way
000008 CC modified and this statement is not removed. Usage by and for commercial
000009 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
000010 CC or send an email to license@isb-sib.ch).
000011 -----
000012 CC EMBL; AF208661; AAF23167.1; -
000013 DR Venom; Neurotoxin; Signal; Amidation;
000014 KW Cleavage on pair of basic residues.
000015 FT SIGNAL 1 26 POTENTIAL.
000016 FT CHAIN 27 57 CONOTOXIN BETX.
000017 FT PROPEP 61 70
000018 FT MOD_RES 57 57
000019 FT AMIDATION (G-58 PROVIDE AMIDE GROUP) (BY
000020 FT SIMILARITY).
000021 SQ SEQUENCE 70 AA; 7900 MW; F6575A2E830AD903 CRC64;

```

```

Query Match      100.0%; Score 49; DB 1; Length 70;
Best Local Similarity 23.5%; Pred. NO. 83;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY      1 XXXXXGCGXXXXXXCX 17
        ::::|:::|:::|:::|:
Db       35 ENDSQCCLNECCWGGCG 51

RESULT      9
YVAU_VACC
ID YVAU_VACC      STANDARD;      PRT;      72 AA.
AC P20530;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 8.8 kDa protein.
DE A ORF U.
OS Vaccinia virus (strain Copenhagen).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OC NCBI_TaxId=10249;
RN [1]
RP SEQUENCE FROM N.A.

```

```

FT SIGNAL 1 18 POTENTIAL.
CH 19 75 HYPOTHETICAL PROTEIN TP0059.
SQ SEQUENCE 75 AA; 8153 MW; E8B9FC3B39633F7D CRC64;

Query Match 100.0%; Score 49; DB 1; Length 75;
Best Local Similarity 23.5%; Pred. NO. 87;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXGCCXXXCCXXCX 17
DB 6 SARSMCCSFSCAKNCA 22
      :::::|:::|:::|:::|:
      6 SARSMCCSFSCAKNCA 22

RESULT 11
HEPC_HUMAN
ID HEPC_HUMAN STANDARD; PRT; 84 AA.
AC P81172;
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Antimicrobial peptide hepcidin precursor (Liver-expressed
DE antimicrobial peptide) (LEAP-1) [Contains: Hepcidin 25 (Hepc25);
DE Hepcidin 20 (Hepc20)].
DE HAMP OR LEAP1 OR HEPC.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RS SEQUENCE FROM N.A., AND SEQUENCE OF 60-84.
RP TISSUE=Liver, and Urine;
RC PubMed=1113131;
RA Park C.H., Valore E.V., Waring A.J., Ganz T.;
RX "Hepcidin: a urinary antimicrobial peptide synthesized in the liver.";
RT J. Biol. Chem. 276:7806-7810(2001).
RN [2]
RS SEQUENCE FROM N.A., SEQUENCE OF 60-84, FUNCTION, TISSUE SPECIFICITY,
RP AND MASS SPECTROMETRY.
RC TISSUE=Liver, and Blood;
RX MEDLINE=20487130; PubMed=11034317;
RA Krause A., Neitz S., Maegert H.J., Schulz A., Forssmann W.G.,
RX Schulz-Knappe P., Adermann K.;
RT "LEAP-1, a novel highly disulfide-bonded human peptide exhibits
RT antimicrobial activity.";
RN FEBS Lett. 480:147-150(2000).
RP [3]
RS SEQUENCE FROM N.A.
RP Lamerdin J.E.;
RA Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: STRONG ANTIMICROBIAL ACTIVITY AGAINST E.COLI ML35P
CC BACTERIUM, N.CINEREA AND WEAKER AGAINST S.EPIDERMIDIS, S.AUREUS
CC ANS GROUP B STREPTOCOCCUS BACTERIA. ACTIVE AGAINST FUNGUS
CC C.ALBICANS. NO ACTIVITY AGAINST P.AERUGINOSA BACTERIUM.
CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN LIVER AND TO A LESSER
CC EXTENT IN HEART AND BRAIN. LOW LEVELS IN LUNG, TONSILS, SALIVARY
CC GLAND, TRACHEA, PROSTATE GLAND, ADRENAL GLAND AND THYROID GLAND.
CC SECRETED INTO THE URINE.
CC -!- MASS SPECTROMETRY: MW=2789.8; METHOD=WALDI; RANGE=60-84.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF309489; AAC239566.1; -
CC DR EMBL; A277280; CAC09419.1; -
CC DR EMBL; AD000684; -; NOT_ANNOTATED_CDS.
CC DR MIM; 606464; -
CC KW Antibiotic; Fungicide; Signal.

```

FT SIGNAL 1 24 POTENTIAL.
FT PROPEP 25 54
FT PEPTIDE 60 84 ANTIMICROBIAL PEPTIDE HEPICIDIN 25.
FT PEPTIDE 65 84 ANTIMICROBIAL PEPTIDE HEPICIDIN 20.
FT DISULFID 66 72 BY SIMILARITY.
FT DISULFID 69 82 BY SIMILARITY.
FT DISULFID 70 81 BY SIMILARITY.
FT DISULFID 73 78 BY SIMILARITY.
SQ SEQUENCE 84 AA; 9408 MW; 5F8DCA23D19D29F7 CRC64;

Query Match 100.0%; Score 49; DB 1; Length 84;
Best Local Similarity 23.5%; Pred. No. 94;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCCXXXCCXXXCX 17
Db 67 IFCCGCCRRSGCMCK 83

RESULT 12
YZX6_HUMAN STANDARD; PRT; 96 AA.
ID YZX6_HUMAN STANDARD; PRT; 96 AA.

AC Q9UF47;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 10.8 kDa protein DAFZP434W0927.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

RC TISSUE-Testis;
RA Bloeker H., Boecher M., Brandt P., Mewes H.-W., Gassenhuber J.,
RA Wilmann S.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----

DR EMBL; AL133660; CAB63773.1;
KW Hypothetical protein.
FT DOMAIN 20 37 CYS-RICH.
FT DOMAIN 23 34 POLY-CYS.
SQ SEQUENCE 96 AA; 10843 MW; 4E0F78EB815E8F79 CRC64;

Query Match 100.0%; Score 49; DB 1; Length 96;
Best Local Similarity 23.5%; Pred. No. 1e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCCXXXCCXXXCX 17
Db 18 TGCYFCCCLCCCNCCC 34

RESULT 13
WAP_CAMDR STANDARD; PRT; 117 AA.
ID WAP_CAMDR STANDARD; PRT; 117 AA.

AC P09837;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Whey acidic protein (WAP).
GN WAP.
OS Camelus dromedarius (Dromedary) (Arabian camel).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.
OX NCBI_TaxID=9838;
RN [1]
RP SEQUENCE.
RC TISSUE-Milk;
RX MEDLINE=86300719; PubMed=3743571;
RA Beg O.U., von Bahr-Lindstrom H., Zaidi Z.H., Joernvall H.;
RT "A camel milk whey protein rich in half-cysteine. Primary structure,
RT assessment of variations, internal repeat patterns, and relationships
RT with neurophysin and other active polypeptides.";
RL Eur. J. Biochem. 159:195-201(1986).
CC -1- FUNCTION: COULD BE A PROTEASE INHIBITOR.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: MILK-SPECIFIC; MAJOR PROTEIN COMPONENT OF MILK
CC WHEY.

CC -1- SIMILARITY: CONTAINS 2 WAP-TYPE DOMAINS.

CC PIR; A24178; A24178.
DR HSPP; O46655; ICJH.
DR InterPro: IPR002221; WAP.
DR Pfam: PF00095; wap; 2.
DR PRINTS; PR00003; 4DISULPHCORE.
DR PRODOM; PD001224; WAP; 1.
DR SMART; SM00217; WAP; 2.
DR PROSITE; PS00317; 4 DISULFIDE CORE; 2.
KW Milk; Whey; Protease inhibitor; Repeat; Phosphorylation.
FT DOMAIN 9 54 WAP 1.
FT DOMAIN 64 111 WAP 2.
FT DISULFID 13 41 BY SIMILARITY.
FT DISULFID 24 46 BY SIMILARITY.
FT DISULFID 28 40 BY SIMILARITY.
FT DISULFID 34 50 BY SIMILARITY.
FT DISULFID 68 99 BY SIMILARITY.
FT DISULFID 80 103 BY SIMILARITY.
FT DISULFID 86 98 BY SIMILARITY.
FT DISULFID 92 107 BY SIMILARITY.
SQ SEQUENCE 117 AA; 12564 MW; 2D9BB6A5A37A921B CRC64;

Query Match 100.0%; Score 49; DB 1; Length 117;
Best Local Similarity 23.5%; Pred. No. 1.2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCCXXXCCXXXCX 17
Db 35 PQGKCCARSPCSRCT 51

RESULT 14
FSHB_BOVIN STANDARD; PRT; 129 AA.
ID FSHB_BOVIN STANDARD; PRT; 129 AA.

AC P04837;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Follicotropin beta chain precursor (follicle-stimulating hormone beta
DE subunit) (FSH-beta) (FSH-B).
GN FSHB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=86313629; PubMed=3092216;
RA Esch F.S., Mason A.J., Cooksey K., Mercado M., Shimasaki S.;
RT "Cloning and DNA sequence analysis of the cDNA for the precursor of
RT the beta chain of bovine follicle stimulating hormone.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:6618-6621(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87053172; PubMed=3096676;

```

DE subunit) (FSH-beta) (FSH-B).
GN FSHB.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=HARTLEY; TISSUE=Pituitary;
RA Suzuki O.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RC
CC -!- FUNCTION: STIMULATES DEVELOPMENT OF FOLLICLE AND SPERMATOGENESIS
CC IN THE REPRODUCTIVE ORGANS.
CC
CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF257212; AAF68975.1; -.
CC InterPro: IPR000359; Cys_knot.
CC InterPro: IPR001545; Glyco_hormone_beta.
CC Pfam: PF00007; Cys_knot; 1.
CC SMART: SM00068; GHb; 1.
CC PROSITE: PS00261; GLYCO_HORMONE_BETA_1; 1.
CC PROSITE: PS00689; GLYCO_HORMONE_BETA_2; 1.
CC Hormone; Glycoprotein; Signal.
CC SIGNAL 1 18 BY SIMILARITY.
CC FT CHAIN 19 129 FOLLITROPIN BETA CHAIN.
CC FT DISULFID 21 69 BY SIMILARITY.
CC FT DISULFID 35 84 BY SIMILARITY.
CC FT DISULFID 38 122 BY SIMILARITY.
CC FT DISULFID 46 100 BY SIMILARITY.
CC FT DISULFID 50 102 BY SIMILARITY.
CC FT DISULFID 105 112 BY SIMILARITY.
CC FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 129 AA; 14694 MW; E543339034DD46A5 CRC64;
SQ
Query Match 100.0%; Score 49; DB 1; Length 129;
Best Local Similarity 23.5%; Pred. No. 1.3e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;
QY 1 XXXXXCCXXXCXXXCX 17
:|||||:|||||:
Db 6 FCFFFCWKAICNGCE 22
RESULT 16
FSHB_HORSE
ID FSHB_HORSE STANDARD; PRT; 129 AA.
AC P01226; Q9TTJ9;
DT 21-JUL-1986 (Rel. 01, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Follitropin beta chain precursor (Follicle-stimulating hormone beta
DE subunit) (FSH-beta) (FSH-B).
GN FSHB.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
[1]
RN SEQUENCE FROM N.A.
RP

```

```

RC TISSUE-Pituitary;
RA Saneyoshi T., Min K., Shiota K.;
RA "Equine follicle-stimulating hormone beta-subunit.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 19-129.
RX MEDLINE=78218213; PubMed=670202;
RA Fujiki Y., Rathnam P., Saxena B.B.;
RT "Amino acid sequence of the beta-subunit of the follicle-stimulating
hormone from equine pituitary glands.";
RL J. Biol. Chem. 253:5363-5368(1978).
CC -1- FUNCTION: STIMULATES DEVELOPMENT OF FOLLICLE AND SPERMATOGENESIS
IN THE REPRODUCTIVE ORGANS.
CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB029157; BAA88560.1; -.
DR PIR; A01494; FTHOB.
DR HSP; P01233; LXUL.
DR InterPro; IPR000359; Cys_knot.
DR Pfam; PF00007; Cys_Knot; 1.
DR SMART; SM00068; GHB; 1.
DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
DR PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
KW Hormone; Glycoprotein; Signal.
FT SIGNAL 1 18
FT CHAIN 19 129 FOLLITROPIN BETA CHAIN.
FT DISULFID 21 69 BY SIMILARITY.
FT DISULFID 35 84 BY SIMILARITY.
FT DISULFID 38 122 BY SIMILARITY.
FT DISULFID 46 100 BY SIMILARITY.
FT DISULFID 50 102 BY SIMILARITY.
FT DISULFID 105 112 BY SIMILARITY.
FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .).
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .).
FT CONFLICT 34 36 ECG -> GCR (IN REF. 2).
FT CONFLICT 40 40 S -> T (IN REF. 2).
FT CONFLICT 64 64 N -> K (IN REF. 2).
FT CONFLICT 99 99 A -> Z (IN REF. 2).
FT CONFLICT 129 129 E -> ZYFVALSY (IN REF. 2).
SQ SEQUENCE 129 AA; 14386 MW; 3916FA71490947CE CRC64;

Query Match 100.0%; Score 49; DB 1; Length 129;
Best Local Similarity 23.5%; pred. No. 1.3e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXCX 17
D6 6 FCFLECCWKAVCNCSCE 22

RESULT 17
FSHB_HUMAN
ID FSHB_HUMAN STANDARD; PRT; 129 AA.
AC P01225;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Follitropin beta chain precursor (Follicle-stimulating hormone beta
subunit) (FSH-beta) (FSH-B).

```

```

GN FSHB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87246070; PubMed=2885163;
RA Watkins P.C., Eddy R., Beck A.K., Vellucci V., Leverone B.,
RA Tanzi R.E., Gusella J.F., Shows T.B.;
RT "DNA sequence and regional assignment of the human
follicle-stimulating hormone beta-subunit gene to the short arm of
human chromosome 11.";
RL DNA 6:205-212(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89174746; PubMed=2494176;
RA Keene J.L., Matzuk M.M., Otani T., Fauser B.C.J.M., Galway A.B.,
RA Hsueh A.J.W., Boime I.;
RT "Expression of biologically active human follitropin in Chinese
hamster ovary cells.";
RL J. Biol. Chem. 264:4769-4775(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=89014411; PubMed=3139991;
RA Jameson J.L., Becker C.B., Lindell C.M., Habener J.F.;
RT "Human follicle-stimulating hormone beta-subunit gene encodes
multiple messenger ribonucleic acids.";
RL Mol. Endocrinol. 2:806-815(1988).
RN [4]
RP SEQUENCE OF 19-129.
RX MEDLINE=89351581; PubMed=3151250;
RA Shome B., Parlow A.F., Liu W.K., Nahm H.S., Wen T., Ward D.N.;
RT "A reevaluation of the amino acid sequence of human follitropin beta-
subunit.";
RL J. Protein Chem. 7:325-339(1988).
RN [5]
RP PRELIMINARY SEQUENCE OF 19-129.
RX MEDLINE=76120602; PubMed=1249074;
RA Saxena B.B., Rathnam P.;
RT "Amino acid sequence of the beta subunit of follicle-stimulating
hormone from human pituitary glands.";
RL J. Biol. Chem. 251:993-1005(1976).
RN [6]
RP PRELIMINARY SEQUENCE OF 19-129.
RX MEDLINE=74262938; PubMed=4835136;
RA Shome B., Parlow A.F.;
RT "Human follicle stimulating hormone: first proposal for the amino
acid sequence of the hormone-specific, beta subunit (hFSHb).";
RL J. Clin. Endocrinol. Metab. 39:203-205(1974).
RN [7]
RP PRELIMINARY ASSIGNMENT OF DISULFIDE BONDS.
RX MEDLINE=81021713; PubMed=6774759;
RA Fujiki Y., Rathnam P., Saxena B.B.;
RT "Studies on the disulfide bonds in human pituitary
follicle-stimulating hormone.";
RL Biochim. Biophys. Acta 624:428-435(1980).
CC -1- FUNCTION: STIMULATES DEVELOPMENT OF FOLLICLE AND SPERMATOGENESIS
IN THE REPRODUCTIVE ORGANS.
CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -1- PHARMACEUTICAL: Available under the names Gonad-F or Metrodin HP
(Serono) and Puregon (Organon).
CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

or send an email to license@isb-sib.ch).

```
CC -----
CC EMBL; M16647; AA52476.1; -.
CC EMBL; M16646; AA52476.1; JOINED.
CC EMBL; M24340; AA52470.1; -.
CC EMBL; M24339; AA52470.1; JOINED.
CC EMBL; M54914; AAB02868.1; -.
CC EMBL; M54913; AAB02868.1; JOINED.
CC PIR; A40920; FTHUB.
CC HSP; P01233; 1XUL.
CC GlycoSuiteDB; P01225; -.
CC MIM; 136530; -.
CC InterPro; IPR00359; Cys_knot.
CC InterPro; IPR001545; Glyco_hormone_beta.
CC Pfam; PF00007; Cys_knot; 1.
CC SMART; SM00068; GHb; 1.
CC PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
CC PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
CC Hormone; Glycoprotein; Pharmaceutical; Signal; Polymorphism.
FT SIGNAL 1 18
FT CHAIN 19 129 FOLLITROPIN BETA CHAIN.
FT DISULFID 21 69 BY SIMILARITY.
FT DISULFID 35 84 BY SIMILARITY.
FT DISULFID 38 122 BY SIMILARITY.
FT DISULFID 46 100 BY SIMILARITY.
FT DISULFID 50 102 BY SIMILARITY.
FT DISULFID 105 112 BY SIMILARITY.
FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .).
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .).
FT VARIANT 20 20 S -> I (IN DBSNP:6170).
FT /FTId=VAR_012047.
SQ SEQUENCE 129 AA; 14700 MW; 902AF00DD7688E2B CRC64;

Query Match 100.0%; Score 49; DB 1; Length 129;
Best Local Similarity 23.5%; Pred. No. 1.3e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 6 FFFLCCKWKAICNSCE 22

RESULT 18
FSHB_PIG
ID FSHB_PIG STANDARD; PRT; 129 AA.
AC P01228;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Follitropin beta chain precursor (Follicle-stimulating hormone beta
DE subunit) (FSH-beta) (FSH-B).
GN FSHB.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LANDRACE-YORKSHIRE; TISSUE=Anterior pituitary;
RX MEDLINE=91063935; PubMed=2174241;
RA Hirai T., Takikawa H., Kato Y.;
RT "The gene for the beta subunit of porcine FSH: absence of consensus
RT oestrogen-responsive element and presence of retroposons.";
RL J. Mol. Endocrinol. 5:147-158(1990).
RN [2]
RP SEQUENCE OF 15-129 FROM N.A.
RC TISSUE=Anterior pituitary;
RX MEDLINE=88196589; PubMed=3129323;
RA Kato Y.;
RT "Cloning and DNA sequence analysis of the cDNA for the precursor of
RT porcine follicle stimulating hormone (FSH) beta subunit.";
RL Mol. Cell. Endocrinol. 55:107-112(1988).
```

```
RN [3]
RP SEQUENCE OF 21-126.
RX MEDLINE=78190610; PubMed=658036;
RA Closset J., Maghuln-Rogister G., Hennen G., Strosberg A.D.;
RT "Porcine follitropin. The amino-acid sequence of the beta subunit.";
RL Eur. J. Biochem. 86:115-120(1978).
CC -!- FUNCTION: STIMULATES DEVELOPMENT OF FOLLICLE AND SPERMATOGENESIS
CC IN THE REPRODUCTIVE ORGANS.
CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/,
CC or send an email to license@isb-sib.ch).
```

```
CC -----
CC EMBL; D00621; BAA00499.1; -.
CC EMBL; M35676; AAA31039.1; -.
CC PIR; A01496; FTPGB.
CC PIR; A48169; A48169.
CC HSP; P01233; 1XUL.
CC InterPro; IPR00359; Cys_knot.
CC InterPro; IPR002400; GF_cysknot.
CC InterPro; IPR001545; Glyco_hormone_beta.
CC Pfam; PF00007; Cys_knot; 1.
CC PRINTS; PR00438; GFCYSKNOT.
CC SMART; SM00068; GHb; 1.
CC PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
CC PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
CC Hormone; Glycoprotein; Signal.
FT SIGNAL 1 20
FT CHAIN 21 129 FOLLITROPIN BETA CHAIN.
FT DISULFID 21 69 BY SIMILARITY.
FT DISULFID 35 84 BY SIMILARITY.
FT DISULFID 38 122 BY SIMILARITY.
FT DISULFID 46 100 BY SIMILARITY.
FT DISULFID 50 102 BY SIMILARITY.
FT DISULFID 105 112 BY SIMILARITY.
FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CONFLICT 32 36 KEECN -> VRCLT (IN REF. 3).
FT CONFLICT 51 52 YT -> TTG (IN REF. 3).
FT CONFLICT 71 72 EK -> YR (IN REF. 3).
FT CONFLICT 125 125 S -> G (IN REF. 3).
SQ SEQUENCE 129 AA; 14605 MW; 20BBCBEDF209E1EA CRC64;

Query Match 100.0%; Score 49; DB 1; Length 129;
Best Local Similarity 23.5%; Pred. No. 1.3e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 6 FFFLCCKWKAICNSCE 22

RESULT 19
FSHB_SHEEP
ID FSHB_SHEEP STANDARD; PRT; 129 AA.
AC P01227;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Follitropin beta chain precursor (Follicle-stimulating hormone beta
DE subunit) (FSH-beta) (FSH-B).
GN FSHB.
OS Ovis aries (Sheep).
```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92029621; PubMed=1930694;
RA Guzman K., Miller C.D., Phillips C.L., Miller W.L.;
RT "The gene encoding ovine follicle-stimulating hormone beta:
RT isolation, characterization, and comparison to a related ovine
RT genomic sequence";
RL DNA Cell Biol. 10:593-601(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89366671; PubMed=2505233;
RA Mountford P.S., Bello P.A., Brandon M.R., Adams T.E.;
RT "Cloning and DNA sequence analysis of the cDNA for the precursor of
RT ovine follicle stimulating hormone beta-subunit";
RL Nucleic Acids Res. 17:6391-6391(1989).
RN [3]
RP SEQUENCE OF 20-129.
RX MEDLINE=82113053; PubMed=6798969;
RA Sairam M.R., Seidah N.G., Chretien M.;
RT "Primary structure of the ovine pituitary follitropin beta-subunit";
RL Biochem. J. 197:541-552(1981).
CC -1- FUNCTION: STIMULATES DEVELOPMENT OF FOLLICLE AND SPERMATOGENESIS
CC IN THE REPRODUCTIVE ORGANS.
CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S64745; AAB20317.1; -;
CC EMBL; X15493; CAA33516.1; -;
CC PIR; A01495; FTSBH.
CC PIR; A40410; A40410.
CC PIR; S05316; S05316.
CC HSP; P01233; IXUL.
CC InterPro; IPR000359; Cys_knot.
CC InterPro; IPR002400; GF_cysknot.
CC Pfam; PF00007; Cys_knot; 1.
CC PRINTS; PR00438; GFCYSKNOT.
CC SMART; SM0068; GHb; 1.
DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
DR PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
KW Hormone; Glycoprotein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 129 FOLLITROPIN BETA CHAIN.
FT DISULFID 21 69 BY SIMILARITY.
FT DISULFID 35 84 BY SIMILARITY.
FT DISULFID 38 122 BY SIMILARITY.
FT DISULFID 46 100 BY SIMILARITY.
FT DISULFID 50 102 BY SIMILARITY.
FT DISULFID 105 112 BY SIMILARITY.
FT CARBOHYD 25 25 N-LINKED (GLCNAC...) (PROBABLE).
FT CARBOHYD 42 42 N-LINKED (GLCNAC...) (PROBABLE).
FT CONFLICT 68 68 A -> T (IN REF. 3).
FT CONFLICT 107 107 R -> T (IN REF. 3).
FT CONFLICT 128 129 RE -> ERZ (IN REF. 3).
SQ SEQUENCE 129 AA; 14669 MW; 83D76DCDC971EF40 CRC64;

Query Match

100.0%; Score 49; DB 1; Length 129;

Best Local Similarity 23.5%; Pred. No. 1.3e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;
QY 1 XXXXXCCXXXCCXXXCX 17
DB 6 FCFLCCWRAICRSC 22
RESULT 20
YJG4_YEAST STANDARD; PRT; 131 AA.
ID YJG4_YEAST
AC F47038;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Hypothetical 13.9 kDa protein in SMC3-MRPL8 intergenic region.
GN YJL064W OR J1120 OR HRC131.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C;
RX MEDLINE=95282514; PubMed=7762302;
RA Vandenbol M., Durand P., Dion C., Portetelle D., Hilger F.;
RT "Sequence of a 17.1 kb DNA fragment from chromosome X of
RT Saccharomyces cerevisiae includes the mitochondrial ribosomal protein
RT L8";
RL Yeast 11:57-60(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z49340; CAA89356.1; -;
CC EMBL; Z34288; CAA84058.1; -;
CC SCD; S0003600; YJL064W.
KW Hypothetical protein.
SQ SEQUENCE 131 AA; 13854 MW; 736E72054593E8F3 CRC64;
Query Match 100.0%; Score 49; DB 1; Length 131;
Best Local Similarity 23.5%; Pred. No. 1.3e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;
QY 1 XXXXXCCXXXCCXXXCX 17
DB 55 GTCCCCCCCCCLCRDSCV 71
RESULT 21
WAP_PIG STANDARD; PRT; 132 AA.
ID WAP_PIG
AC Q4655; O97559;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Whey acidic protein precursor (WAP).
GN WAP.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 20-56 AND 93-132.
RX MEDLINE=98174264; PubMed=9513079;
RA Simpson K.J., Bird P., Shaw D., Nicholas K.R.;
RT "Molecular characterisation and hormone-dependent expression of the


```
RA Elleman T.C., Dopheide T.A.;
RT "The sequence of SCMK-B2b, a high-sulfur protein from wool keratin.";
RJ J. Biol. Chem. 247:3909-3909(1972).
CC -1- FUNCTION: THE KERATIN PRODUCTS OF MAMMALIAN EPIDERMAL DERIVATIVES
CC SUCH AS WOOL AND HAIR CONSIST OF MICROFIBRILS EMBEDDED IN A RIGID
CC MATRIX OF OTHER PROTEINS. THE MATRIX PROTEINS INCLUDE THE HIGH-
CC SULFUR AND HIGH-TYROSINE KERATINS, HAVING MOLECULAR WEIGHTS OF
CC 6-20 kDa, WHEREAS THE MICROFIBRILS CONTAIN THE LARGER, LOW-SULFUR
CC KERATINS (40-56 kDa).
CC -1- MISCELLANEOUS: THE SOURCE OF THIS KERATIN IS LINCOLN WOOL.
DR PIR; A02838; KRSHB.
DR InterPro: IPR002494; Keratin_B2.
DR Pfam: PF01500; Keratin_B2; 1.
KW Keratin; Acetylation; Repeat.
FT MOD_RES 1 1 ACETYLATION.
FT REPEAT 26 35
FT REPEAT 36 45
FT REPEAT 46 55
FT REPEAT 56 65
FT VARIAT 77 77 D -> G (IN MINOR COMPONENT).
SQ SEQUENCE 156 AA; 16040 MW; EE549ADCFF1FB39 CRC64;

Query Match 100.0%; Score 49; DB 1; Length 156;
Best Local Similarity 23.5%; Pred. NO. 1.4e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 121 YQAASCCRPSCGQSCC 137

RESULT 24
KRUH HUMAN STANDARD; PRT; 169 AA.
AC P26371;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE keratin, ultra high-sulfur matrix protein A (UHS keratin A) (UHS
DE Kera).
GN KRN1 OR UHSK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Follicle;
RX MEDLINE=91115951; PubMed=1703541;
RA McKinnon P.J., Powell B.C., Rogers G.E.;
RT "Structure and expression of genes for a class of cysteine-rich
RT proteins of the cuticle layers of differentiating wool and hair
RT follicles.";
RL J. Cell Biol. 111:2587-2600(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99148005; PubMed=10023043;
RA Perez C., Auriol J., Gerst C., Bernard B.A., Egly J.-M.;
RT "Genomic organization and promoter characterization of two human UHS
RT keratin genes.";
RL Gene 227:137-148(1999).
CC -1- FUNCTION: THE KERATIN PRODUCTS OF MAMMALIAN EPIDERMAL DERIVATIVES
CC SUCH AS WOOL AND HAIR CONSIST OF MICROFIBRILS EMBEDDED IN A RIGID
CC MATRIX OF OTHER PROTEINS. THE MATRIX PROTEINS INCLUDE THE HIGH-
CC SULFUR AND HIGH-TYROSINE KERATINS, HAVING MOLECULAR WEIGHTS OF
CC 6-20 kDa, WHEREAS THE MICROFIBRILS CONTAIN THE LARGER, LOW-SULFUR
CC KERATINS (40-56 kDa).
CC -1- TISSUE SPECIFICITY: CUTICLE LAYERS OF DIFFERENTIATING
CC HAIR FOLLICLES.
CC -1- DEVELOPMENTAL STAGE: AT A LATE STAGE OF FIBER DIFFERENTIATION.
CC -1- DOMAIN: MAINLY COMPOSED OF CYS-RICH (CR), GLY-RICH (GR) AND SER-
CC RICH (SR) REPEATS.
```

```
CC -1- SIMILARITY: BELONGS TO THE UHS KERATIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X55293; CAA39005.1; -.
DR EMBL; AJ006693; CAA07189.1; -.
DR HSP; P04355; 2MRT.
DR MIM; 148021; -.
KW Keratin; Repeat; Multigene family.
SQ SEQUENCE 169 AA; 16276 MW; 219B14FEEB49D4AB CRC64;

Query Match 100.0%; Score 49; DB 1; Length 169;
Best Local Similarity 23.5%; Pred. NO. 1.5e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 77 CSQCSCCKPCCSSGCG 93

RESULT 25
KR2A SHEEP STANDARD; PRT; 171 AA.
AC P02438;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-AUG-1991 (Rel. 19, Last annotation update)
DE Keratin, high-sulfur matrix protein, B2A.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83299218; PubMed=6193483;
RA Powell B.C., Sleight M.J., Ward K.A., Rogers G.E.;
RT "Mammalian keratin gene families: organisation of genes coding for
RT the B2 high-sulfur proteins of sheep wool.";
RL Nucleic Acids Res. 11:5327-5346(1983).
RN [2]
RP SEQUENCE.
RX MEDLINE=73224964; PubMed=4679226;
RA Elleman T.C.;
RT "The amino acid sequence of protein SCMK-B2a from the high-sulphur
RT fraction of wool keratin.";
RL Biochem. J. 130:833-845(1972).
CC -1- FUNCTION: THE KERATIN PRODUCTS OF MAMMALIAN EPIDERMAL DERIVATIVES
CC SUCH AS WOOL AND HAIR CONSIST OF MICROFIBRILS EMBEDDED IN A RIGID
CC MATRIX OF OTHER PROTEINS. THE MATRIX PROTEINS INCLUDE THE HIGH-
CC SULFUR AND HIGH-TYROSINE KERATINS, HAVING MOLECULAR WEIGHTS OF
CC 6-20 kDa, WHEREAS THE MICROFIBRILS CONTAIN THE LARGER, LOW-SULFUR
CC KERATINS (40-56 kDa).
CC -1- MISCELLANEOUS: THE SOURCE OF THIS KERATIN IS LINCOLN WOOL.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X01610; CAA25757.1; -.
DR PIR; A02837; KRSHHA.
DR PIR; S07910; S07910.
```

DR InterPro: IPR002494; Keratin_B2.
 DR Pfam: PF01500; Keratin_B2; 1.
 KW Keratin; Acetylation; Repeat.

FT INIT_MET 0 0 ACETYLATION.
 FT MOD_RES 1 1
 FT REPEAT 26 35
 FT REPEAT 36 45
 FT REPEAT 46 55
 FT REPEAT 56 65
 FT REPEAT 66 75
 FT REPEAT 66 75

FT VARIANT 12 12 I -> T (IN MINOR COMPONENT).
 FT VARIANT 22 22 S -> N (IN MINOR COMPONENT).
 FT VARIANT 23 23 P -> F (IN MINOR COMPONENT).
 SQ SEQUENCE 171 AA; 17603 MW; 79290B74A1B9044C CRC64;

Query Match 100.0%; Score 49; DB 1; Length 171;
 Best Local Similarity 23.5%; Pred. No. 1.5e+02;
 Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Oy 1 XXXXXCCXXCCXXCCX 17
 Db 131 YQAASCCRPSCGSGCC 147

RESULT 26

DR KR2D_SHEEP STANDARD; PRT; 181 AA.
 AC P08131;

DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 01-AUG-1991 (Rel. 19, Last annotation update)
 DE Keratin, high-sulfur matrix protein, B2D.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=83299218; PubMed=6193483;
 RA Powell B.C., Sleight M.J., Ward K.A., Rogers G.E.;
 RT "Mammalian keratin gene families: organisation of genes coding for
 the B2 high-sulphur protein of sheep wool.";

RL Nucleic Acids Res. 11:5327-5346(1983).
 CC -!- FUNCTION: THE KERATIN PRODUCTS OF MAMMALIAN EPIDERMAL DERIVATIVES
 CC SUCH AS WOOL AND HAIR CONSIST OF MICROFIBRILS EMBEDDED IN A RIGID
 CC MATRIX OF OTHER PROTEINS. THE MATRIX PROTEINS INCLUDE THE HIGH-
 CC SULFUR AND HIGH-TYROSINE KERATINS, HAVING MOLECULAR WEIGHTS OF
 CC 6-20 kDa, WHEREAS THE MICROFIBRILS CONTAIN THE LARGER, LOW-SULFUR
 CC KERATINS (40-56 kDa).
 CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

CC EMBL: X01610; CAA25759.1; -.
 DR PIR: S07911; KRSHHD.

DR InterPro: IPR002494; Keratin_B2.
 DR Pfam: PF01500; Keratin_B2; 1.

KW Keratin; Repeat.

FT INIT_MET 0 0
 FT DOMAIN 26 85 6 X 10 AA TANDEM REPEATS.
 FT REPEAT 26 35 1.
 FT REPEAT 36 45 2.
 FT REPEAT 46 55 3.
 FT REPEAT 56 65 4.
 FT REPEAT 66 75 5.
 FT REPEAT 76 85 6.

SQ SEQUENCE 181 AA; 18679 MW; D3E4874E21757B12 CRC64;

Query Match 100.0%; Score 49; DB 1; Length 181;
 Best Local Similarity 23.5%; Pred. No. 1.6e+02;
 Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Oy 1 XXXXXCCXXCCXXCCX 17
 Db 141 YQAASCCRPSCGSGCC 157

RESULT 27

DR KRUB_HUMAN

ID KRUB_HUMAN STANDARD; PRT; 194 AA.
 AC Q75690;

DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Keratin, ultra high-sulfur matrix protein B (UHS keratin B) (UHS
 DE KerB).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=99148005; PubMed=10023043;
 RA Perez C., Aurioi J., Gerst C., Bernard B.A., Egly J.-M.;
 RT "Genomic organization and promoter characterization of two human UHS
 RT keratin genes.";
 RL Gene 227:137-148(1999).
 CC -!- FUNCTION: THE KERATIN PRODUCTS OF MAMMALIAN EPIDERMAL DERIVATIVES
 CC SUCH AS WOOL AND HAIR CONSIST OF MICROFIBRILS EMBEDDED IN A RIGID
 CC MATRIX OF OTHER PROTEINS. THE MATRIX PROTEINS INCLUDE THE HIGH-
 CC SULFUR AND HIGH-TYROSINE KERATINS, HAVING MOLECULAR WEIGHTS OF
 CC 6-20 kDa, WHEREAS THE MICROFIBRILS CONTAIN THE LARGER, LOW-SULFUR
 CC KERATINS (40-56 kDa).
 CC

CC -!- TISSUE SPECIFICITY: CUTICLE LAYERS OF DIFFERENTIATING
 CC HAIR FOLLICLES.
 CC

CC -!- DOMAIN: MAINLY COMPOSED OF CYS-RICH (CR), GLY-RICH (GR) AND SER-
 CC RICH (SR) REPEATS.
 CC -!- SIMILARITY: BELONGS TO THE UHS KERATIN FAMILY.
 CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

CC EMBL: AJ006692; CAA07188.1; -.
 DR HSSP: P04355; 2MPT.

KW Keratin; Repeat; Multigene family.

SQ SEQUENCE 194 AA; 18094 MW; 8607B2AE906FE44A CRC64;

Query Match 100.0%; Score 49; DB 1; Length 194;
 Best Local Similarity 23.5%; Pred. No. 1.7e+02;
 Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Oy 1 XXXXXCCXXCCXXCCX 17
 Db 96 SCYKPCCGSGGSGCC 112

RESULT 28

DR CSP_TORCA

ID CSP_TORCA STANDARD; PRT; 195 AA.
 AC P56101;

DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)

Proc. Natl. Acad. Sci. U.S.A. 89:3503-3507(1992).

```

[2]
RN SEQUENCE FROM N.A.
RP TISSUE-Bone marrow, and Lung;
RC TISSUE-Bone marrow, and Lung;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
[3]
RN TOPOLGY.
RX MEDLINE=94171760; PubMed=7510285;
RA Marken J.S., Bajorath J., Edwards C.P., Farr A.G., Schlieven G.L.,
RA Hellstrom I., Hellstrom K.E., Aruffo A.;
RT "Membrane topology of the L6 antigen and identification of the protein
RT epitope recognized by the L6 monoclonal antibody.";
RL J. Biol. Chem. 269:7397-7401(1994).
CC -1- SUBUNIT: PRESENT IN HIGH MOLECULAR WEIGHT COMPLEXES IN TUMOR
CC CELLS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED ON LUNG, BREAST, COLON, AND
CC OVARIAN CARCINOMAS. IT IS ALSO PRESENT ON SOME NORMAL CELLS,
CC ENDOTHELIAL CELLS IN PARTICULAR.
CC -1- SIMILARITY: BELONGS TO THE TETRASPANIN (TM4SF) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M90657; AAA36158.1; -.
DR EMBL; BC008442; AAH08442.1; -.
DR EMBL; BC010166; AAH10166.1; -.
DR PIR; A42926; A42926.
DR MIM; 191155; -.
KW Glycoprotein; Antigen; Transmembrane.
FT DOMAIN 1 9 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 10 30 PROBABLE.
FT DOMAIN 31 49 EXTRACELLULAR (PROBABLE).
FT TRANSMEM 50 70 PROBABLE.
FT DOMAIN 71 93 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 94 114 PROBABLE.
FT DOMAIN 115 161 EXTRACELLULAR (PROBABLE).
FT TRANSMEM 162 182 PROBABLE.
FT DOMAIN 183 202 CYTOPLASMIC (PROBABLE).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 159 159 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 202 AA; 21632 MW; 80D82C47902B8CD4 CRC64;
```

Query Match 100.0%; Score 49; DB 1; Length 202;
Best Local Similarity 23.5%; Pred. No. 1.7e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

```

Oy 1 XXXXXCCXXXCCXXXCC 17
    ::::|:::|:::|:::|
Db 73 DCCGCCGCGHNGKRC A 89
```

RESULT 35
RGSH_CHICK STANDARD; PRT; 210 AA.

```

AC Q9PWA0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Regulator of G-protein signaling 17 (RGS17).
GN RGS17.
OS Gallus gallus (Chicken).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
CC Gallus.
OX NCBI_TaxID=9031;
```

```

[1]
RN SEQUENCE FROM N.A.
RP TISSUE-Dorsal root ganglion;
RX MEDLINE=95348269; PubMed=10419452;
RA Jordan J.D., Carey K.D., Stork P.J.S., Iyengar R.;
RT "Modulation of Rap activity by direct interaction of Galphao with Rap1
RT GTPase-activating protein.";
RL J. Biol. Chem. 274:21507-21510(1999).
CC -1- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC THEIR INACTIVE GDP-BOUND FORM.
CC -1- SUBCELLULAR LOCATION: Membrane-bound (Potential).
CC -1- PTM: FATTY ACYLATED. HEAVILY PALMITOYLATED IN THE CYSTEINE STRING
CC MOTIF (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 RGS DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF151968; AAD45948.1; -.
DR HSSP; P49795; ICMZ.
DR InterPro; IPR000342; RGS.
DR Pfam; PF00615; RGS; 1.
DR PRINTS; PR01301; RGS/PROTEIN.
DR PRODOM; PD001580; RGS; 1.
DR SMART; SM00315; RGS; 1.
DR PROSITE; PS00132; RGS; 1.
KW Signal transduction inhibitor; Membrane; Lipoprotein; Palmitate.
FT DOMAIN 28 40 POLY-CYS.
FT DOMAIN 84 200 RGS.
SQ SEQUENCE 210 AA; 24326 MW; 6581AAD5BADDEE7C CRC64;
```

Query Match 100.0%; Score 49; DB 1; Length 210;
Best Local Similarity 23.5%; Pred. No. 1.7e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

```

Oy 1 XXXXXCCXXXCCXXXCC 17
    ::::|:::|:::|:::|
Db 23 RPNTCCFCWCSCSCS 39
```

RESULT 36
RGSH_HUMAN STANDARD; PRT; 210 AA.

```

AC Q9UGC6; Q9UJS8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Regulator of G-protein signaling 17 (RGS17).
GN RGS17.
OS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
```

SEQUENCE FROM N.A.
Phillimore B.;
Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

RP TISSUE-Brain;
RC TISSUE-Brain;
RA Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

```

-1- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC THEIR INACTIVE GDP-BOUND FORM.
CC -1- SUBCELLULAR LOCATION: Membrane-bound (Potential).
```


DB 34 PSRNPCLCWCSCCS 50

RESULT 40
RGSJ_HUMAN
ID RGSJ_HUMAN STANDARD: PRT: 217 AA.
AC P49795;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Regulator of G-protein signaling 19 (RGS19) (G-alpha interacting
protein) (GAIP protein).
GN RGS19 OR GAIP OR GNAI3IP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96102226; PubMed=8524874;
RA de Vries L., Mousli M., Wurmser A., Farquhar M.G.;
RT "GAIP, a protein that specifically interacts with the trimeric G
protein alpha 13, is a member of a protein family with a highly
conserved core domain."
RL Proc. Natl. Acad. Sci. U.S.A. 92:11916-11920(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggeley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,
RA Buck D., Burrill W., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark L.N., Clark S.V., Clee C.M.,
RA Coulson A., Coville G.J., Deadman R., Dhali P., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grahm D.V., Griffiths C., Griffiths M.N.D., Guilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Lehtvasilho M.H., Leversha M., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McLeay J.C., Nickerson T.,
RA Milne S., Mistry D., Moore M.J.F., Mullikin J.C., Peck A.I.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Ramsay H.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Showkeen R., Sims S.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Steward C.A., Sulston J.E.,
RA Skuce C.D., Smith M.L., Soderlund C., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Swann M., Sycamore N., Taylor R., Vaudin M., Wall M., Wallis J.M.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20."
RL Nature 414:865-871(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP PALMITOYLATION.
RX MEDLINE=97140307; PubMed=8986788;
RA de Vries L., Elenko E., Hubler L., Jones T.L.Z., Farquhar M.G.;
RT "GAIP is membrane-anchored by palmitoylation and interacts with the
activated (GTP-bound) form of G alpha i subunits."
RL Proc. Natl. Acad. Sci. U.S.A. 93:15203-15208(1996).
RN [5]
RP STRUCTURE BY NMR OF 79-206.
RX MEDLINE=99384138; PubMed=10452897;
RA de Alba E., De Vries L., Farquhar M.G., Tjandra N.;
RT "Solution structure of human GAIP (Galpha interacting protein): a
regulator of G protein signaling."
RL J. Mol. Biol. 291:927-939(1999).

[6]
RN INHIBITION.
RP MEDLINE=98421527; PubMed=9748280;
RA Wang J., Ducret A., Tu Y., Kozasa T., Aebersold R., Ross E.M.;
RT "RGS21, a Gz-selective RGS protein in brain: Structure, membrane
association, regulation by G(alpha)z phosphorylation, and
RT relationship to a Gz GAPase-activating protein subfamily.";
RL J. Biol. Chem. 273:26014-26025(1998).
RN [7]
RP PHOSPHORYLATION OF SER-151, AND MUTAGENESIS OF SER-151.
RX MEDLINE=20564368; PubMed=10993892;
RA Ogier-Denis E., Pattinre S., El Benna J., Codogno P.;
RT "Erk1/2-dependent phosphorylation of Galpha-interacting protein
RT stimulates its GTPase accelerating activity and autophagy in human
colon cancer cells.";
RL J. Biol. Chem. 275:39090-39095(2000).
CC -!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC ACTIVITY OF G PROTEIN ALPHA SUBUNIT THEREBY DRIVING THEM INTO
CC THEIR INACTIVE GDP-BOUND FORM. BINDS TO G-ALPHA SUBFAMILY 1
CC MEMBERS, WITH THE ORDER G(I)A3 > G(I)A1 > G(O)A >> G(2)A/G(I)A2.
CC ACTIVITY ON G(2)-ALPHA IS INHIBITED BY PHOSPHORYLATION AND
CC PALMITOYLATION OF THE G-PROTEIN.
CC -!- SUBUNIT: INTERACTS WITH GIPC PDZ DOMAIN.
CC -!- SUBCELLULAR LOCATION: Membrane-bound.
CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN LUNG, PLACENTA, LIVER
CC AND HEART ALSO EXPRESS HIGH LEVELS OF GAIP.
CC -!- PTM: FATTY ACYLATED. HEAVILY PALMITOYLATED IN THE CYSTEINE STRING
CC MOTIF.
CC -!- PTM: PHOSPHORYLATED, MAINLY ON SERINE RESIDUES.
CC -!- SIMILARITY: CONTAINS 1 RGS DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X91809; CAA62919.1; -;
DR EMBL; AL590548; CAD11902.1; -;
DR EMBL; BC001318; AAH01318.1; -;
DR PDB; 1CMZ; 10-NOV-99.
DR MIM; 605071; -;
DR InterPro; IPR000342; RGS.
DR Pfam; PF00615; RGS; 1.
DR PRINTS; PD01301; RGS-PROTEIN.
DR ProDom; PD001580; RGS; 1.
DR SMART; SM00315; RGS; 1.
DR PROSITE; PS01332; RGS; 1.
DR Signal transduction inhibitor; Membrane; Lipoprotein; Palmitate;
KW Phosphorylation; Autophagy; 3D-structure.
FT DOMAIN 90 206 RGS.
FT DOMAIN 39 49 POLY-CYS.
FT DOMAIN 207 217 INTERACTS WITH GIPC.
FT MOD_RES 24 24 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 151 151 PHOSPHORYLATION (BY ERK1/2).
FT MUTAGEN 151 151 S->A: DIMINISHES GAP ACTIVITY TOWARDS
FT G(I)-ALPHA3 AND AUTOPHAGY IN COLON CANCER
FT CELLS.
SQ SEQUENCE 217 AA; 24635 MW; 925A5687DC222CDB CRC64;

Query Match 100.0%; Score 49; DB 1; Length 217;
Best Local Similarity 23.5%; Pred. No. 1.8e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCXXXXXXCXCX 17
Db 34 PSRNPCLCWCSCCS 50
RESULT 41

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A., AND VARIANT ARG-32.
 RC STRAIN=129/B6, AND BALB/C;
 RP Barker S.A., Wang J., Ross E.M.;
 RT "A mouse ortholog of RGS21";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN SEQUENCE OF 5-239 FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Hippocampus;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Konno H., Adachi J., Fukuda S.,
 RA Arakawa T., Hara A., Fukunishi Y., Kondo S., Yamanaka I.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Norodone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT Functional annotation of a full-length mouse cDNA collection.*;
 RL Nature 409:685-690(2001).
 CC -1- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
 CC ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
 CC THEIR INACTIVE GDP-BOUND FORM. BINDS SELECTIVELY TO G(2)-ALPHA AND
 CC IS INHIBITED BY THE PHOSPHORYLATION AND PALMITOYLATION OF THE G-
 CC SUBCELLULAR LOCATION: Membrane-bound (Potential).
 CC -1- SUBCELLULAR LOCATION: Membrane-bound (Potential).
 CC -1- PTM: FATTY ACYLATED. HEAVILY PALMITOYLATED IN THE CYSTEINE STRING
 CC MOTIF (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 RGS DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercia
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL: AF191554; AAF05757.1; -;
 DR EMBL: AF191552; AAF05756.1; -;
 DR EMBL: AK013773; BAB28987.1; -;
 DR HSP: P49795; 1CMZ.
 DR MGD: MGI:1929866; Rgs20.
 DR InterPro: IPR000342; RGS.
 DR Pfam: PF00615; RGS.1.
 DR PRINTS: PR01301; RGS-PROTEIN.
 DR ProDom: PD001580; RGS; 1.
 DR SMART: SM00315; RGS; 1.
 DR PROSITE: PS00132; RGS; 1.
 KW Signal transduction inhibitor; Membrane; Lipoprotein; Palmitate;
 KW Polymorphism. 59 71 POLY-CYS.
 FT DOMAIN 113 229 RGS.
 FT VARIANT 32 32 M -> R (IN BALB/C).
 SQ SEQUENCE 239 AA; 26986 MW; F383923163A44D18 CRC64;

CC AND CSP3/CSP29; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN WIDE RANGE OF SYNAPTIC TERMINALS:
CC EMBRYONIC NERVOUS SYSTEM, LARVAL NEUROMUSCULAR JUNCTIONS, ADULT
CC VISUAL SYSTEM (NEUROPILO OF OPTIC GANGLIA AND TERMINAL OF RI-8
CC PHOTORECEPTORS) AND THORACIC NEUROMUSCULAR JUNCTIONS. ALSO
CC EXPRESSED IN NON-NEURONAL CELLS: FOLLICLE CELLS, SPERMATHECA,
CC TESTIS AND EJACULATORY BULB. LOW LEVEL OF EXPRESSION IS FOUND IN
CC MANY NEURONAL AND NON-NEURONAL TISSUES.
CC -1- PTM: FATTY ACYLATED. HEAVILY PALMITOYLATED IN THE CYSTEINE STRING
CC MOTIF.
CC -1- SIMILARITY: CONTAINS 1 J DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M63421; AAA28432.1; -;
CC EMBL; M63008; AAA28431.1; -;
CC EMBL; AF0571167; AAD09428.1; -;
CC EMBL; AF0571167; AAD09430.1; -;
CC EMBL; AF0571167; AAD09431.1; -;
CC EMBL; AE003597; AAF51816.1; -;
CC EMBL; AE003597; AAF51817.1; -;
CC HSSP; P25685; IHDJ.
CC FlyBase; FBgn0004179; Csp.
CC InterPro; IPR001623; DnaJ.N.
CC Pfam; PF00226; DnaJ.1.
CC PRINTS; PR00625; DNAJPROTEIN.
CC SMART; SM00271; DnaJ.1.
CC PROSITE; PS00636; DNAJ_1; 1.
CC PROSITE; PS00766; DNAJ_2; 1.
CC Membrane; Lipoprotein; Palmitate; Alternative splicing.
FT DOMAIN 15 84
FT DOMAIN 121 131 POLY-CYS.
FT VARSPIC 154 174 MISSING (IN ISOFORM CSP3 AND ISOFORM
FT CSP2).
FT VARSPIC 243 249 DMVNOKY -> GI (IN ISOFORM CSP3).
FT VARIANT 71 71 N -> D (IN STRAIN BERKELEY).
SQ SEQUENCE 249 AA; 26896 MW; 3EF97C3BF2553EB8 CRC64;

Query Match 100.0%; Score 49; DB 1; Length 249;
Best Local Similarity 23.5%; Pred. NO. 2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCCXXXCCXXXCX 17
Db 116 AVITGCCCCGCCGCCCN 132

RESULT 45
ID EIO_NPVAC STANDARD; PRT; 261 AA.
AC P41710;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Immediate-early protein IE-0.
GN IE-0.
OS Autographa californica nuclear polyhedrosis virus (ACMNPV).
CC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
CC Nucleopolyhedrovirus.
CC NCBI_TaxID=46015;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C6;
RX MEDLINE=94303173; PubMed=8030224;
RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
RT "The complete DNA sequence of Autographa californica nuclear

polyhedrosis virus.";
RL Virology 202:586-605(1994).
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L22858; AAA66771.1; -;
CC InterPro; IPR001841; Znf_ring.
CC PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
CC PROSITE; PS00089; ZF_RING_2; 1.
CC Early protein; Zinc-finger.
FT ZN_FING 212 257 RING-TYPE.
SQ SEQUENCE 261 AA; 30109 MW; 7721E0C528EC3CBE CRC64;

Query Match 100.0%; Score 49; DB 1; Length 261;
Best Local Similarity 23.5%; Pred. NO. 2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCCXXXCCXXXCX 17
Db 225 LKPECCCEATGACCV 241

RESULT 46
ID APHC_HUMAN STANDARD; PRT; 267 AA.
AC Q9NUN7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alkaline phosphatase (EC 3.5.1.-) (APHC) (Alkaline phosphatase).
GN APHC.
OS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC TISSUE=Kidney;
RX MEDLINE=21336649; PubMed=11356846;
RA Mao C., Xu R., Szulc Z.M., Bielawska A., Galadari S.H., Obeid L.M.;
RT "Cloning and characterization of a novel human alkaline phosphatase. A
RT mammalian enzyme that hydrolyzes phosphoceramides.";
RL J. Biol. Chem. 276:26577-26588(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H., Sugawara M.,
RA Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Hydrolyzes only phosphoceramides into phosphoglycerol and
CC free fatty acid. Does not have reverse activity.
CC -1- ENZYME REGULATION: Activated by Ca(2+) and inhibited by Zn(2+).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum and Golgi apparatus.
CC -1- TISSUE SPECIFICITY: Ubiquitously expressed. Highest expression in
CC placenta.
CC -1- SIMILARITY: BELONGS TO THE ALKALINE PHOSPHATASE FAMILY.
CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
CC FRAMESHIFTS.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF214454; AK071923.1; -
DR EMBL; AK002100; BA92085.1; ALT_FRAME.
KW Hydrolase; Transmembrane.
FT TRANSMEM 29 51 POTENTIAL.
FT TRANSMEM 63 84 POTENTIAL.
FT TRANSMEM 94 111 POTENTIAL.
FT TRANSMEM 118 137 POTENTIAL.
FT TRANSMEM 147 169 POTENTIAL.
FT TRANSMEM 174 196 POTENTIAL.
FT TRANSMEM 216 238 POTENTIAL.
FT CARBOHYD 24 24 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 267 AA; 31565 MW; 936A8C600D9F38E1 CRC64;

Query Match 100.0%; Score 49; DB 1; Length 267;
Best Local Similarity 23.5%; Pred. No. 2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 95 PMYSCCIFVYCMPECF 111
:|||||:|||||:

RESULT 47
APHC_MOUSE STANDARD; PRT; 267 AA.
ID APHC_MOUSE Q9D0X4; Q9D3J4;
AC Q9D099; Q9D0X4; Q9D3J4;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Alkaline phytoeceramidase (EC 3.5.1.-) (APHC) (Alkaline ceramidase).
GN APHC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo, and Head;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King C., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Wynshaw-Boris A., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Hayashizaki Y.;
RA "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- FUNCTION: Hydrolyzes only phytoceramide into phytosphingosine and
CC free fatty acid. Does not have reverse activity (By similarity).
CC -!- ENZYME REGULATION: Activated by Ca(2+) and inhibited by Zn(2+) (By
CC similarity).

```

```

CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum and Golgi apparatus (By similarity).
CC -!- SIMILARITY: BELONGS TO THE ALKALINE CERAMIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AK011668; BAB27768.1; -
DR EMBL; AK004287; BAB23250.1; -
DR EMBL; AK017361; BAB30708.1; -
KW Hydrolase; Transmembrane.
FT TRANSMEM 30 52 POTENTIAL.
FT TRANSMEM 65 84 POTENTIAL.
FT TRANSMEM 94 111 POTENTIAL.
FT TRANSMEM 118 137 POTENTIAL.
FT TRANSMEM 147 169 POTENTIAL.
FT TRANSMEM 174 196 POTENTIAL.
FT TRANSMEM 216 238 POTENTIAL.
FT CONFLICT 110 110 C -> F (IN REF. 1; BAB23250).
FT CONFLICT 262 262 E -> D (IN REF. 1; BAB30708).
SQ SEQUENCE 267 AA; 31564 MW; 4EF7E912CBB4E9BB CRC64;

Query Match 100.0%; Score 49; DB 1; Length 267;
Best Local Similarity 23.5%; Pred. No. 2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 95 PMYSCCIFVYCMPECF 111
:|||||:|||||:

RESULT 48
Y856_TREPA STANDARD; PRT; 325 AA.
ID Y856_TREPA
AC O83828;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Hypothetical protein TP0856 precursor.
GN TP0856.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NICHOLS;
RX MEDLINE=98332770; PubMed=965876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodgren E., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RA "Complete genome sequence of Treponema pallidum, the syphilis
RA spirochete.";
RL Science 281:375-388(1998).
CC -!- SIMILARITY: BELONGS TO THE UPF0164 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```


CC IS INHIBITED BY PHOSPHORYLATION AND PALMITOYLATION OF THE G-
CC PROTEIN.
CC -1- SUBCELLULAR LOCATION: Membrane-bound.
CC -1- ALTERNATIVE PRODUCTS: 6 ISOFORMS: 1, 2, 3, 4, 5 AND 6 (SHOWN
CC HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: ISOFORM 5 IS SPECIFICALLY EXPRESSED IN BRAIN
CC WITH HIGH LEVELS IN THE CAUDATE NUCLEUS AND TEMPORAL LOBE.
CC -1- PTM: FATTY ACYLATED. HEAVILY PALMITOYLATED IN THE CYSTEINE STRING
CC MOTIF (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 RGS DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF060877; AAC62009.2; -
DR EMBL; AF366054; AAK54122.1; -
DR EMBL; AF366055; AAK54123.1; -
DR EMBL; AF366056; AAK54124.1; -
DR EMBL; AF366057; AAK54125.1; -
DR EMBL; AF074979; AAC62013.1; -
DR EMBL; AY046538; AAL03971.1; -
DR HSP; P49795; ICMZ.
DR InterPro: IPR000342; RGS.
DR Pfam: PF00615; RGS; 1.
DR PRINTS; PR01301; RGS-PROTEIN.
DR PRODOM; PD001580; RGS; 1.
DR SMART; SM00315; RGS; 1.
DR PROSITE; PS00132; RGS; 1.
KW Signal transduction inhibitor; Membrane; Lipoprotein; Palmitate;
KW Alternative splicing.
FT DOMAIN 208 220
FT DOMAIN 262 378
FT VARSPLIC 1 55
FT
FT
FT
FT VARSPLIC 56 170
FT
FT VARSPLIC 1 208
FT VARSPLIC 1 236
FT VARSPLIC 209 220
FT
FT VARSPLIC 237 247
FT
FT VARSPLIC 3 55
FT SEQUENCE 388 AA; 43691 MW; F44796D271F1765F CRC64;
SQ

Query Match 100.0%; Score 49; DB 1; Length 388;
Best Local Similarity 23.5%; Pred. No. 2.6e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXCXXCXX 17
Db 203 RGNACCCFCWCCSCS 219

Search completed: July 1, 2002, 12:29:10
Job time: 122 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 12:26:43 ; Search time 25.29 seconds
(without alignments)
116.288 Million cell updates/sec

Title: US-09-493-795A-1
Perfect score: 49
Sequence: 1 XXXXXCXXXXXXCX 17

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : SPTREMBL_19:.*
1: sp_archaea:.*
2: sp_bacteria:.*
3: sp_fungi:.*
4: sp_human:.*
5: sp_invertebrate:.*
6: sp_mammal:.*
7: sp_mbc:.*
8: sp_organelle:.*
9: sp_phage:.*
10: sp_plant:.*
11: sp_rodent:.*
12: sp_virus:.*
13: sp_vertebrate:.*
14: sp_unclassified:.*
15: sp_virus:.*
16: sp_bacteriap:.*
17: sp_archaeap:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 49 | 100.0 | 29 | 5 Q9TWM7 | Q9Twm7 conus glori |
| 2 | 49 | 100.0 | 36 | 5 Q9BP43 | Q9bp43 conus penna |
| 3 | 49 | 100.0 | 46 | 4 Q16861 | Q16861 homo sapien |
| 4 | 49 | 100.0 | 48 | 2 Q9RFP5 | Q9rfp5 mycoplasma |
| 5 | 49 | 100.0 | 48 | 5 Q9BP44 | Q9bp44 conus arena |
| 6 | 49 | 100.0 | 50 | 4 Q9H279 | Q9h279 homo sapien |
| 7 | 49 | 100.0 | 60 | 12 Q85002 | Q85002 porcine res |
| 8 | 49 | 100.0 | 62 | 5 Q9U619 | Q9u619 conus imper |
| 9 | 49 | 100.0 | 63 | 5 Q9N301 | Q9n301 caenorhabdi |
| 10 | 49 | 100.0 | 63 | 5 Q9NLG5 | Q9nlg5 leishmania |
| 11 | 49 | 100.0 | 67 | 5 Q9SS57 | Q9ss57 drosophila |
| 12 | 49 | 100.0 | 68 | 5 Q9BPJ7 | Q9bpj7 conus texti |
| 13 | 49 | 100.0 | 68 | 5 Q9BP11 | Q9bp11 conus penna |
| 14 | 49 | 100.0 | 69 | 5 Q9BPH7 | Q9bph7 conus ventr |
| 15 | 49 | 100.0 | 70 | 5 Q9BH51 | Q9bh51 conus tessu |
| 16 | 49 | 100.0 | 70 | 5 Q9U323 | Q9u323 conus betul |

```

17 49 100.0 70 5 Q9BPJ6  Q9bpj6 conus texti
18 49 100.0 70 5 Q9BPH8  Q9bph8 conus ventr
19 49 100.0 70 5 Q9BP45  Q9bp45 conus penna
20 49 100.0 70 5 Q9BP44  Q9bp44 conus tessu
21 49 100.0 71 5 Q9BP10  Q9bp10 conus texti
22 49 100.0 71 5 Q9BPH6  Q9bph6 conus ventr
23 49 100.0 72 5 Q9BPJ5  Q9bpj5 conus texti
24 49 100.0 72 5 Q9BPH9  Q9bph9 conus tessu
25 49 100.0 73 5 Q9NLG6  Q9nlg6 leishmania
26 49 100.0 76 5 Q9BP43  Q9bp43 conus ventr
27 49 100.0 76 5 Q9BPC3  Q9bpc3 conus ventr
28 49 100.0 76 10 Q40855  Q40855 picea glauc
29 49 100.0 77 5 Q9BPC2  Q9bpc2 conus ventr
30 49 100.0 78 2 Q9X4X8  Q9x4x8 pseudomonas
31 49 100.0 83 11 Q9JM45  Q9jma5 mus musculu
32 49 100.0 83 11 Q9EQ21  Q9eq21 mus musculu
33 49 100.0 84 4 Q9BY68  Q9by68 homo sapien
34 49 100.0 84 11 Q9BMH3  Q9bmh3 rattus norv
35 49 100.0 85 4 Q96S60  Q96s60 homo sapien
36 49 100.0 85 16 Q9PGB1  Q9pgb1 xylella fas
37 49 100.0 87 12 Q913Y8  Q913y8 avian hepat
38 49 100.0 90 5 Q9N4F0  Q9n4f0 caenorhabdi
39 49 100.0 90 5 Q9BPJ4  Q9bpj4 conus texti
40 49 100.0 98 4 Q9BYR4  Q9byr4 homo sapien
41 49 100.0 101 11 Q9ERH9  Q9erh9 cavia porce
42 49 100.0 105 4 Q9BYP8  Q9byp8 homo sapien
43 49 100.0 106 4 Q9BYU7  Q9byu7 homo sapien
44 49 100.0 109 4 Q9BRJ8  Q9brj8 homo sapien
45 49 100.0 109 11 Q9D227  Q9d227 mus musculu
46 49 100.0 111 4 Q9BYQ1  Q9byq1 homo sapien
47 49 100.0 112 4 Q14634  Q14634 homo sapien
48 49 100.0 113 11 Q9D228  Q9d228 mus musculu
49 49 100.0 114 4 Q9BYQ9  Q9byq9 homo sapien
50 49 100.0 120 2 Q9RNT5  Q9rnt5 solar lake

```

ALIGNMENTS

RESULT 1

```

Q9TWM7  PRELIMINARY;      PRT;   29 AA.
AC  Q9TWM7:
DT  01-MAY-2000 (TRENBLrel. 13, Created)
DT  01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DE  01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE  DELTA-CONOTOXIN GMVIA.
OS  Conus gloriamaris (Glory of the sea).
OC  Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC  Neogastropoda; Conoidea; Conidae; Conus.
OX  NCBI_TaxID=37336;
RN  [1]
RP  SEQUENCE.
RX  MEDLINE=95001845; PubMed=7918355;
RA  Shon K.J., Hasson A., Spira M.E., Cruz L.J., Gray W.R., Olivera B.M.;
RT  "Delta-conotoxin GmVIA, a novel peptide from the venom of Conus
RT  gloriamaris."
RL  Biochemistry 33:11420-11425(1994).
SQ  SEQUENCE 29 AA; 3360 MW; 9F61A72DC6615D97 CRC64;

```

Query Match 100.0%; Score 49; DB 5; Length 29;
Best Local Similarity 23.5%; Pred. No. 56;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCXXXXXXCX 17

Db 13 PIFQNCRCGWCNCLFCV 29

RESULT 2

```

Q9BP43  PRELIMINARY;      PRT;   36 AA.
ID  Q9BP43

```

```
AC Q9BP43;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CONOTOXIN SCAFFOLD III/IV (FRAGMENT).
OS Conus pennaceus (feathered cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=37335;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21105969; PubMed=11158371;
RA Conticello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,
RA Fainzilber M.;
RT "Mechanisms for Evolving Hypervariability: The Case of Conopeptides.";
RL Mol. Biol. Evol. 18:120-131(2001).
DR EMBL; AF215130; AAG60535.1; -
DT NON_TER 1
FT SEQUENCE 36 AA; 3683 MW; 04C74E40FF71141C CRC64;
SQ SEQUENCE 36 AA; 3683 MW; 04C74E40FF71141C CRC64;

Query Match 100.0%; Score 49; DB 5; Length 36;
Best Local Similarity 23.5%; Pred. No. 65;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
   ::::|:::|:::|:::|:::|
DB 16 KGQPLCCPFGGCHLCH 32

RESULT 3
Q16861
ID Q16861 PRELIMINARY; PRT; 46 AA.
AC Q16861;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE SUPER CYSTEINE RICH PROTEIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lemasson I., Devaux C., Mesnard J.M.;
RT "Characterization of a novel human cysteine repeat-containing cDNA
   encoding a domain with 23 successive cysteines.";
DR Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
EMBL; U63332; AAB05810.1; -
DT NON_TER 1
FT SEQUENCE 46 AA; 5061 MW; 311922FE40A44E8F CRC64;
SQ SEQUENCE 46 AA; 5061 MW; 311922FE40A44E8F CRC64;

Query Match 100.0%; Score 49; DB 4; Length 46;
Best Local Similarity 23.5%; Pred. No. 77;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
   ::::|:::|:::|:::|:::|
DB 6 PDRSRCCCCCCCCCCCC 22

RESULT 4
Q9RFP5
ID Q9RFP5 PRELIMINARY; PRT; 48 AA.
AC Q9RFP5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE ORF2
OS Mycoplasma fermentans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
```

```
OX NCBI_TaxID=2115;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PG18;
RX MEDLINE=99115554; PubMed=9916088;
RA Calcutt M.J., Kim M.F., Karpas A.B., Muhlradt P.F., Wise K.S.;
RT "Differential posttranslational processing confers intraspecies
   variation of a major surface lipoprotein and a macrophage-activating
   lipopeptide of Mycoplasma fermentans.";
RT Infect. Immun. 67:760-771(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PG18;
RX MEDLINE=20069643; PubMed=10601219;
RA Calcutt M.J., Lavrarr J.L., Wise K.S.;
RT "IS1630 of Mycoplasma fermentans, a novel IS30-type insertion element
   that targets and duplicates inverted repeats of variable length and
   sequence during insertion.";
RL J. Bacteriol. 181:7597-7607(1999).
DR EMBL; AF179376; AAF15567.1; -
SQ SEQUENCE 48 AA; 5319 MW; 668836FA3592B2C7 CRC64;

Query Match 100.0%; Score 49; DB 2; Length 48;
Best Local Similarity 23.5%; Pred. No. 80;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
   ::::|:::|:::|:::|:::|
DB 22 CTTKDCCKNSCCSCCK 38

RESULT 5
Q9BP44
ID Q9BP44 PRELIMINARY; PRT; 48 AA.
AC Q9BP44;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DSC-2001 (TrEMBLrel. 19, Last annotation update)
DE CONOTOXIN SCAFFOLD III/IV (FRAGMENT).
OS Conus arenatus.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=89451;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21105969; PubMed=11158371;
RA Conticello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,
RA Fainzilber M.;
RT "Mechanisms for Evolving Hypervariability: The Case of Conopeptides.";
RL Mol. Biol. Evol. 18:120-131(2001).
DR EMBL; AF215129; AAG60534.1; -
DT NON_TER 1
FT SEQUENCE 48 AA; 5563 MW; F830DCF58A7E747F CRC64;
SQ SEQUENCE 48 AA; 5563 MW; F830DCF58A7E747F CRC64;

Query Match 100.0%; Score 49; DB 5; Length 48;
Best Local Similarity 23.5%; Pred. No. 80;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
   ::::|:::|:::|:::|:::|
DB 25 LRQECVCVPFCGGCV 41

RESULT 6
Q9H2T9
ID Q9H2T9 PRELIMINARY; PRT; 50 AA.
AC Q9H2T9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE G ALPHA INTERACTING PROTEIN (FRAGMENT).
```

| | | |
|----|--|---|
| GN | GAIP. | |
| OS | Homo sapiens (Human). | |
| OC | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | |
| OC | NCBI_TaxID=9606; | |
| RP | [1] | |
| RP | SEQUENCE FROM N.A. | |
| RX | MEDLINE=20545389; PubMed=11090272; | |
| RT | Ito E., Xie, Gx, Maruyama K., Palmer P. P.; | |
| RT | "A core-promoter region functions bi-directionally for human opioid- | |
| RT | receptor-like gene ORL1 and its 5'-adjacent gene GAIP."; | |
| RL | J. Mol. Biol. 304:259-270(2000). | |
| DR | ENBL; AF217688; AAG42352.1; -. | |
| DR | NON_TER | 50 |
| DR | SEQUENCE | 50 AA; 5324 MW; 09DA78A00BE18A43 CRC64; |
| FT | | |

```

RESULT      7
Q85002      PRELIMINARY;      PRT;      60 AA.
ID          Q85002;
AC          01-NOV-1996 (TReMBLrel. 01, Created)
DT          01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DE          01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE          SPIKE PROTEIN (FRAGMENT).
OS          Porcine respiratory coronavirus.
OC          Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC          Coronaviridae; Coronavirus.
CX          NCBI_TaxID=11146;
RN          [1]
RP          SEQUENCE FROM N.A.
RC          STRAIN=86/137004;
RX          MEDLINE=91170940; PubMed=1848593;
RA          Page K.W., Mawdlitt K.L., Britton P.:
RT          "Sequence comparison of the 5'-end of mRNA 3 from transmissible
RT          gastroenteritis virus and porcine respiratory coronavirus.";
RL          J. Gen. Virol. 72:579-587(1991).
DR          EMBL; D00658; BAA00548.1; -.
DR          InterPro: IPR002552; Corona_S2.
DR          Pfam; PF01601; Corona_S2; 1.
DR          NON_TER
;          1
;          SEQUENCE 60 AA; 6811 MW; D4D2FA8386D4686 CRC64;

```

| | | |
|--------|--|---|
| RESULT | 8 | |
| Q9U619 | | |
| ID | Q9U619 | PRELIMINARY; PRT; 62 AA. |
| AC | Q9U619; | |
| DT | 01-MAY-2000 | (TrEMBLrel. 13, Created) |
| DF | 01-MAY-2000 | (TrEMBLrel. 13, Last sequence update) |
| DT | 01-DEC-2001 | (TrEMBLrel. 19, Last annotation update) |
| DE | CONTOTOXIN IMIIA. | |
| GN | IMIIA. | |
| OS | Conus imperialis (Imperial cone). | |
| OC | Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda; | |

OC Neogastropoda: Conoidea; Conidae; Conus.
OX NCBI_taxid=35631;
[1]
RN SEQUENCE FROM N. A.
RP Zhao D., Huang P.
RA "Conus imperialis conotoxin IMiia precursor mRNA."
RT Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF200595; AAF12824.1;
DR EMBL: AF200595; AAF12824.1;
SQ SEQUENCE 62 AA: 6830 MW: BF0D811758C3047D CRC64;

```

RESULT      9
QN301      PRELIMINARY;      PRT;      63 AA.
ID          QN301
AC          QN301;
DT          01-OCT-2000 (TrEMBLrel. 15, Created)
DT          01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT          01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE          DE HP0CTHECAL 6.4 KDA PROTEIN.
DE          Y65B4BL.6.
OS          Caenorhabditis elegans.
OS          Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditoidea;
OC          Rhabditidae; Peloderinae; Caenorhabditis.
OX          NCBI_TaxID=6239;
RN          [1]
RP          SEQUENCE FROM N.A.
RC          STRAIN-BRISTOL N2;
RX          MEDLINE=99069613; PubMed=9851916;
RA          None;
RT          Genome sequence of the nematode C. elegans: a platform for
RT          investigating biology. The C. elegans Sequencing Consortium.
RL          Science 282:2012-2018(1998).

```

| | |
|--------|---|
| RESULT | 10 |
| Q9NLG5 | |
| ID | Q9NLG5 |
| AC | Q9NLG5; |
| DT | 01-OCT-2000 (TrEMBLrel. 15, Created) |
| DT | 01-OCT-2000 (TrEMBLrel. 15, Last sequence update) |
| | PRELIMINARY; |
| | PRT; 63 AA. |

```
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PROBABLE CTG26 ALTERNATE OPEN READING FRAME (FRAGMENT).
GN LM26.477
OS Leishmania major.
OC Eukaryota; Euzoenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIDLIN;
RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL160493; CAB98095.1; -.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001211; PLP_A2.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS00118; PA2_His; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 63 AA; 6680 MW; 976F09B837D61713 CRC64;

Query Match 100.0%; Score 49; DB 5; Length 63;
Best Local Similarity 29.4%; Pred. No. 96;
Matches 5; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
   :|::|::|::|::|::|:
Db 17 TXCCCCCYHGCCCCY 33

RESULT 11
ID Q95S57 PRELIMINARY; PRT; 67 AA.
AC Q95S57;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GM08588P.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY060947; AAL28495.1; -.
SQ SEQUENCE 67 AA; 7533 MW; 58654467C32A8805 CRC64;

Query Match 100.0%; Score 49; DB 5; Length 67;
Best Local Similarity 23.5%; Pred. No. 1e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
   :|::|::|::|::|::|:
Db 33 GCWLGCLIPCCIDDCM 49

RESULT 12
ID Q9BPJ7 PRELIMINARY; PRT; 68 AA.
AC Q9BPJ7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CONOTOXIN SCAFFOLD III/IV.
OS Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OX NCBI_TaxID=117992;
RN [1]
RP SEQUENCE FROM N.A.
RL MEDLINE=21105969; PubMed=11158371;
RA Conticello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,
RA Fainzilber M.;
SQ SEQUENCE 68 AA; 7603 MW; 54CA322FFA7CB1B0 CRC64;
```

```
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6494;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21105969; PubMed=11158371;
RA Conticello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,
RA Fainzilber M.;
RT "Mechanisms for Evolving Hypervariability: The Case of Conopeptides.";
RL Mol. Biol. Evol. 18:120-131(2001).
DR EMBL; AF214926; AAG60354.1; -.
SQ SEQUENCE 68 AA; 7603 MW; 54CA322FFA7CB1B0 CRC64;

Query Match 100.0%; Score 49; DB 5; Length 68;
Best Local Similarity 23.5%; Pred. No. 1e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
   :|::|::|::|::|::|:
Db 48 DATRCCNAGFCRFGCT 64

RESULT 13
ID Q9BPI1 PRELIMINARY; PRT; 68 AA.
AC Q9BPI1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CONOTOXIN SCAFFOLD III/IV.
OS Conus pennaceus (Feathered cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=37335;
RN [1]
RP SEQUENCE FROM N.A.
RL MEDLINE=21105969; PubMed=11158371;
RA Conticello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,
RA Fainzilber M.;
RT "Mechanisms for Evolving Hypervariability: The Case of Conopeptides.";
RL Mol. Biol. Evol. 18:120-131(2001).
DR EMBL; AF214943; AAG60371.1; -.
DR InterPro; IPR001230; Prenyltn.
DR PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
SQ SEQUENCE 68 AA; 7643 MW; A239484F7D277249 CRC64;

Query Match 100.0%; Score 49; DB 5; Length 68;
Best Local Similarity 23.5%; Pred. No. 1e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
   :|::|::|::|::|::|:
Db 48 LGQRVCCPESCTDRCL 64

RESULT 14
ID Q9BPH7 PRELIMINARY; PRT; 69 AA.
AC Q9BPH7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CONOTOXIN SCAFFOLD III/IV.
OS Conus ventricosus (Mediterranean cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=117992;
RN [1]
RP SEQUENCE FROM N.A.
RL MEDLINE=21105969; PubMed=11158371;
RA Conticello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,
RA Fainzilber M.;
SQ SEQUENCE 69 AA; 7643 MW; A239484F7D277249 CRC64;
```

RT "Mechanisms for Evolving Hypervariability: The Case of Conopeptides."
RL Mol. Biol. Evol. 18:120-131(2001).
DR EMBL: AF214948; AAG60376.1; -
SQ SEQUENCE 69 AA; 7936 MW; 739F45A9818BF68D CRC64;

Query Match 100.0%; Score 49; DB 5; Length 69;
Best Local Similarity 23.5%; Pred. No. 1e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 50 PREQECPEQWCDGGCD 66

RESULT 15
Q9BH51 ID Q9BH51 PRELIMINARY; PRT; 70 AA.
AC Q9BH51;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CONOTOXIN SCAFFOLD III/IV PRECURSOR.
OS Conus tessulatus.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=101317;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21105969; PubMed=11158371;
RA Conticello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,
RA Fainzilber M.;
RT "Mechanisms for Evolving Hypervariability: The Case of Conopeptides."
RL Mol. Biol. Evol. 18:120-131(2001).
DR EMBL: AF215091; AAG60512.1; -
DR EMBL: AF214946; AAG60374.1; -
SQ SEQUENCE 70 AA; 8061 MW; BC6AD9C7263DEB7D CRC64;

Query Match 100.0%; Score 49; DB 5; Length 70;
Best Local Similarity 23.5%; Pred. No. 1e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 51 LGORCCDWQWCDGACD 67

RESULT 16
Q9U323 ID Q9U323 PRELIMINARY; PRT; 70 AA.
AC Q9U323;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE BETX TOXIN.
GN BETX.
OS Conus betulinus (Beech cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=89764;
RN [1]
RP SEQUENCE FROM N.A.
RA Fan C.X.;
RT "A conotoxin with novel cysteine framework from Conus betulinus of south China sea."
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF208661; AAF23167.1; -
FT CHAIN 27 57 BETX TOXIN.
SQ SEQUENCE 70 AA; 7900 MW; F6575A2E830AD903 CRC64;

Query Match 100.0%; Score 49; DB 5; Length 70;

Best Local Similarity 23.5%; Pred. No. 1e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 35 ENDSQCCLNECCWGGCG 51

RESULT 17
Q9BPJ6 ID Q9BPJ6 PRELIMINARY; PRT; 70 AA.
AC Q9BPJ6;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CONOTOXIN SCAFFOLD III/IV.
OS Conus textile (Cloth-of-gold cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6494;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21105969; PubMed=11158371;
RA Conticello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,
RA Fainzilber M.;
RT "Mechanisms for Evolving Hypervariability: The Case of Conopeptides."
RL Mol. Biol. Evol. 18:120-131(2001).
DR EMBL: AF214927; AAG60355.1; -
SQ SEQUENCE 70 AA; 7848 MW; 92825D82F3BA9B68 CRC64;

Query Match 100.0%; Score 49; DB 5; Length 70;
Best Local Similarity 23.5%; Pred. No. 1e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 49 DPNKRCDDSECSYSCW 65

RESULT 18
Q9BP8 ID Q9BP8 PRELIMINARY; PRT; 70 AA.
AC Q9BP8;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CONOTOXIN SCAFFOLD III/IV.
OS Conus ventricosus (Mediterranean cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=117992;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21105969; PubMed=11158371;
RA Conticello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,
RA Fainzilber M.;
RT "Mechanisms for Evolving Hypervariability: The Case of Conopeptides."
RL Mol. Biol. Evol. 18:120-131(2001).
DR EMBL: AF214947; AAG60375.1; -
SQ SEQUENCE 70 AA; 7990 MW; 4FB31E9C477DB326 CRC64;

Query Match 100.0%; Score 49; DB 5; Length 70;
Best Local Similarity 23.5%; Pred. No. 1e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 50 PRQRCCEPDWCDGCD 66

RESULT 19

RX MEDLINE=211105969; PubMed=11158371;
 RA Conticello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,
 RA Fainzilber M.;
 RA "Mechanisms for Evolving Hypervariability: The Case of Conopeptides.";
 RT Mol. Biol. Evol. 18:120-131(2001).
 Z EMBL; AF214950; AAG60378.1 -
 Z SEQUENCE 70 AA; 7998 MW; 4D865A3403EFBDC1 CRC64;

| | |
|--------|---|
| RESULT | 22 |
| Q9BPH6 | |
| ID | PRELIMINARY; |
| AC | PRT; 71 AA. |
| DT | (TREMBLrel. 17, Created) |
| DT | (TREMBLrel. 17, Last sequence update) |
| DT | (TREMBLrel. 19, Last annotation update) |

OS *Conus ventricosus* (Mediterranean *Conus*).
 OS Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.
 OX NCBI_TaxID=117992;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=211105969; PubMed=11158371.
 RX Conticello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,
 RA Fainzilber M.,
 RA "Mechanisms for Evolving Hypervariability: The Case of Conopeptides.",
 RT Mol. Biol. Evol. 18:120-131(2001).
 RL EMBL; AF214949; AAG60377.1.
 SR SEQUENCE 71 AA; 8107 mw; 7194A31364123B98 CRC64;
 SQ

| | |
|--------|---|
| RESULT | 23 |
| Q9BPJ5 | |
| ID | PRELIMINARY; PRT; 72 AA. |
| AC | Q9BPJ5; |
| DT | 01-JUN-2001 (TRENBLrel. 17, Created) |
| DT | 01-JUN-2001 (TRENBLrel. 17, Last sequence update) |
| DT | 01-JUN-2001 (TRENBLrel. 17, Last annotation update) |
| DT | 01-DEC-2001 (TRENBLrel. 19, Last annotation update) |

OS Conus textile (Cloth-of-gold cone).
 OT Eukaryota; Metazoa; Mollusca; Gastropoda;
 PA Caenogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.
 OX NCBI_TaxID=6494;
 RI {1}
 RN SEQUENCE FROM N.A.
 RP MEDLINE=21105969; PubMed=11158371.
 RX Conticello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,
 RA Fainzilber M.;
 RA "Mechanisms for Evolving Hypervariability: The Case of Conopeptides."; Mol. Biol. Evol. 18:120-131(2001).
 RL

DR EMBL; AF214928; AAG60356.1; -
SQ SEQUENCE 72 AA; 7820 MW; EE680CB65BB43320 CRC64;

Query Match 100.0%; Score 49; DB 5; Length 72;
Best Local Similarity 23.5%; Pred. No. 1.1e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
|:::|:::|:::|:::|:::|
Db 52 KKKQCCPPVACMGCE 68

RESULT 24

Q9BPH9 ID Q9BPH9 PRELIMINARY; PRT; 72 AA.
AC Q9BPH9; 72 AA.
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CONOTOXIN SCAFFOLD III/IV.
OS Conus tessulatus.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=101317;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21105969; PubMed=11158371;
RA Conticello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,
RA Fainzilber M.;
RT "Mechanisms for Evolving Hypervariability: The Case of Conopeptides.";
RL Mol. Biol. Evol. 18:120-131(2001).
DR EMBL; AF214945; AAG60373.1; -
SQ SEQUENCE 72 AA; 8149 MW; F8A582CAD89EC2E7 CRC64;

Query Match 100.0%; Score 49; DB 5; Length 72;
Best Local Similarity 23.5%; Pred. No. 1.1e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
|:::|:::|:::|:::|:::|
Db 51 LGORDCCVMPWCDGACD 67

RESULT 25

Q9NLG6 ID Q9NLG6 PRELIMINARY; PRT; 73 AA.
AC Q9NLG6; 73 AA.
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE PROBABLE CYSTEINE-RICH PROTEIN (FRAGMENT).
GN LM26.476.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-FRIEDLIN;
RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL160493; CAB98094.1; -
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001211; PDP_A2.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS00118; PA2_HIS; UNKNOWN_2.
FT NON_TER
SQ SEQUENCE 73 AA; 7630 MW; 6E624D5AD40B834E CRC64;

Query Match 100.0%; Score 49; DB 5; Length 73;
Best Local Similarity 29.4%; Pred. No. 1.1e+02;

Matches 5; Conservative 12; Mismatches 0; Indels 0; Gaps 0;
QY 1 XXXXXCCXXXCCXXXCX 17
|:::|:::|:::|:::|:::|
Db 50 XGCCCCCVHGCCCCCY 66

RESULT 26

Q9BPC4 ID Q9BPC4 PRELIMINARY; PRT; 76 AA.
AC Q9BPC4; 76 AA.
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CONOTOXIN SCAFFOLD VI/VII.
OS Conus ventricosus (Mediterranean cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=117992;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21105969; PubMed=11158371;
RA Conticello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,
RA Fainzilber M.;
RT "Mechanisms for Evolving Hypervariability: The Case of Conopeptides.";
RL Mol. Biol. Evol. 18:120-131(2001).
DR EMBL; AF215010; AAG60438.1; -
SQ SEQUENCE 76 AA; 8534 MW; CBD1381A05B64D82 CRC64;

Query Match 100.0%; Score 49; DB 5; Length 76;
Best Local Similarity 23.5%; Pred. No. 1.1e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
|:::|:::|:::|:::|:::|
Db 59 WNNPPCCSWTCDYCK 75

RESULT 27

Q9BPC3 ID Q9BPC3 PRELIMINARY; PRT; 76 AA.
AC Q9BPC3; 76 AA.
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CONOTOXIN SCAFFOLD VI/VII.
OS Conus ventricosus (Mediterranean cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=117992;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21105969; PubMed=11158371;
RA Conticello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,
RA Fainzilber M.;
RT "Mechanisms for Evolving Hypervariability: The Case of Conopeptides.";
RL Mol. Biol. Evol. 18:120-131(2001).
DR EMBL; AF215011; AAG60439.1; -
SQ SEQUENCE 76 AA; 8658 MW; E3E8A8A05B64EB2 CRC64;

Query Match 100.0%; Score 49; DB 5; Length 76;
Best Local Similarity 23.5%; Pred. No. 1.1e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
|:::|:::|:::|:::|:::|
Db 59 WNNPPCCSWTCDYCK 75

RESULT 28

Q40855

DR EMBL; AF038507; AAF28736.1; -;
DR EMBL; AK008437; BAB25668.1; -;
DR EMBL; AK002502; BAB22150.1; -;
DR MGD; MGI:1915605; 0610010M13R1k.
DR MGD; MGI:1922859; 2010204K13R1k.
SQ SEQUENCE 83 AA; 8909 MW; 275B78E46428A3DE CRC64;

Query Match 100.0%; Score 49; DB 11; Length 83;
Best Local Similarity 23.5%; Pred. No. 1.2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

OY 1 XXXXXCXXXXXXCXXC 17
Db 6 KPAKSCCLPCCCCYCR 22
:::|||||:::|:

RESULT 32

ID Q9EQ21 PRELIMINARY; PRT; 83 AA.
AC Q9EQ21;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PROHEPCIDIN.
GN HAMP OR HEPC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21269329; PubMed=11113132;
RA Pigeon C.; Ilyin G., Courselaud B., Leroyer P., Turlin B., Brissot P.,
RA Loreal O.;
RT "A new mouse liver-specific gene, encoding a protein homologous to
RT human antimicrobial peptide hepcidin, is overexpressed during iron
RT overload";
RL J. Biol. Chem. 276:7811-7819(2001).
DR EMBL; AF297664; AAG49293.1; -;
DR MGD; MGI:1933533; Hamp.
SQ SEQUENCE 83 AA; 9352 MW; C76423EA46260B18 CRC64;

Query Match 100.0%; Score 49; DB 11; Length 83;
Best Local Similarity 23.5%; Pred. No. 1.2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

OY 1 XXXXXCXXXXXXCXXC 17
Db 66 IFCCCKCNSQCIGCK 82
:::|||||:::|:

RESULT 33

ID Q9BY68 PRELIMINARY; PRT; 84 AA.
AC Q9BY68;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE PUTATIVE LIVER TUMOR REGRESSOR.
GN PLTR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Jung J.-W., Shin W.-S., Yoon Y., Lee S.-T.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF131292; AAK14912.1; -;
SQ SEQUENCE 84 AA; 9438 MW; E694CA23D18EB441 CRC64;

Query Match 100.0%; Score 49; DB 4; Length 84;
Best Local Similarity 23.5%; Pred. No. 1.2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

OY 1 XXXXXCXXXXXXCXXC 17
Db 67 IFCCGCHRSKCGMCK 83
:::|||||:::|:

RESULT 34

ID Q99MH3 PRELIMINARY; PRT; 84 AA.
AC Q99MH3;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PROHEPCIDIN.
GN HEPC.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=SPRAGUE DAWLEY; TISSUE=LIVER;
RX MEDLINE=21269329; PubMed=11113132;
RA Pigeon C.; Ilyin G., Courselaud B., Leroyer P., Turlin B., Brissot P.,
RA Loreal O.;
RT "A new mouse liver-specific gene, encoding a protein homologous to
RT human antimicrobial peptide hepcidin, is overexpressed during iron
RT overload";
RL J. Biol. Chem. 276:7811-7819(2001).
DR EMBL; AF344185; AAK12966.1; -;
DR InterPro; IPR001230; Prenyltn.
DR PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
SQ SEQUENCE 84 AA; 9286 MW; 65E0D5FE4F44A0B0 CRC64;

Query Match 100.0%; Score 49; DB 11; Length 84;
Best Local Similarity 23.5%; Pred. No. 1.2e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

OY 1 XXXXXCXXXXXXCXXC 17
Db 67 LFCCCKCKNSQGLCCI 83
:::|||||:::|:

RESULT 35

ID Q96S60 PRELIMINARY; PRT; 85 AA.
AC Q96S60;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE KERATIN ASSOCIATED PROTEIN 1.7.
GN HKAP1.7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Aoki N., Shimomura Y., Langbein L., Rogers M.A., Schweizer J., Ito M.;
RT "Two Novel Human Keratin-Associated Proteins Expressed in the Cortical
RT Layer of the Hair Follicle Have been Termed hKAP1.6 and hKAP1.7";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB055057; BAB61031.1; -;
SQ SEQUENCE 85 AA; 9109 MW; DA45842FE23DCBD4 CRC64;

Query Match 100.0%; Score 49; DB 4; Length 85;
Best Local Similarity 23.5%; Pred. No. 1.2e+02;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17

Db 56 HAEASCCRPSCGQSCC 72

RESULT 36

Q9PGB1 ID Q9PGB1 PRELIMINARY; PRT; 85 AA.

AC Q9PGB1; MEDLINE=20365717; PubMed=10910347;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE HYPOTHETICAL PROTEIN XF0391.

GN XF0391.

OS Xylella fastidiosa.

CC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;

NCBI_TaxID=2371;

NCBI_TaxID=2371;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=9A5C;

RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,

RA Alvares R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,

RA Barros M.P., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,

RA Bueno M.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,

RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,

RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,

RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,

RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,

RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,

RA Ho P.L., Honeisel J.D., Junqueira M.L., Kemper E.L., Kicajima J.P.,

RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,

RA Lenos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,

RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,

RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,

RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,

RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,

RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,

RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,

RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,

RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,

RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,

RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,

RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,

RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,

RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,

RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;

RL "The genome sequence of the plant pathogen Xylella fastidiosa.";

RL Nature 406:151-159(2000).

DR EMBL; AE003890; AAF83201.1;

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 85 AA; 9217 MW; 8A7A565B2225D0D60 CRC64;

Query Match 100.0%; Score 49; DB 16; Length 85;

Best Local Similarity 23.5%; Pred. No. 1.2e+02;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17

Db 3 CCAVLCCAVLCCAVLCC 19

RESULT 37

Q913Y8 ID Q913Y8 PRELIMINARY; PRT; 87 AA.

AC Q913Y8; MEDLINE=20365717; PubMed=10910347;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DE HYPOTHETICAL PROTEIN XF0391.

GN XF0391.

OS Xylella fastidiosa.

CC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;

NCBI_TaxID=2371;

NCBI_TaxID=2371;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=9A5C;

RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,

RA Alvares R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,

RA Barros M.P., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,

RA Bueno M.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,

RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,

RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,

RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,

RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,

RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,

RA Ho P.L., Honeisel J.D., Junqueira M.L., Kemper E.L., Kicajima J.P.,

RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,

RA Lenos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,

RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,

RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,

RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,

RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,

RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,

RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,

RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,

RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,

RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,

RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,

RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,

RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,

RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,

RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;

RL "The genome sequence of the plant pathogen Xylella fastidiosa.";

RL Nature 406:151-159(2000).

DR EMBL; AE003890; AAF83201.1;

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 85 AA; 9217 MW; 8A7A565B2225D0D60 CRC64;

DE HYPOTHETICAL 9.1 KDA PROTEIN.

OS Avian hepatitis E virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage.

OX NCBI_TaxID=172851;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21446678; PubMed=11562538;

RA Hagshenas G., Shivaprasad H.L., Woolcock P.R., Read D.H., Meng X.J.;

RT "Genetic identification and characterization of a novel virus related

RT to human hepatitis E virus from chickens with hepatitis-splenomegaly

RT syndrome in the United States.";

RL J. Gen. Virol. 82:2449-2462(2001).

DR EMBL; AY043166; AAL13368.1;

RN Hypothetical protein.

SQ SEQUENCE 87 AA; 9121 MW; 329711739BE48F4D CRC64;

Query Match 100.0%; Score 49; DB 12; Length 87;

Best Local Similarity 23.5%; Pred. No. 1.2e+02;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17

Db 19 CRCVDCSCLCQAAGCQ 35

RESULT 38

Q9N4F0 ID Q9N4F0 PRELIMINARY; PRT; 90 AA.

AC Q9N4F0; MEDLINE=99069613; PubMed=9851916;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE HYPOTHETICAL 10.5 KDA PROTEIN.

GN Y71H2B.4.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BRISTOL N2;

RX MEDLINE=99069613; PubMed=9851916;

RA None;

RT "Genome sequence of the nematode C. elegans: a platform for

RT investigating biology. The C. elegans Sequencing Consortium.";

RL Science 282:2012-2018(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=BRISTOL N2;

RA Leonard S., Spalding L.;

RT "The sequence of C. elegans cosmid Y71H2B.";

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=BRISTOL N2;

RA Waterston R.;

RT "Direct Submission.";

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC024202; AAF36033.1;

KW Hypothetical protein.

SQ SEQUENCE 90 AA; 10525 MW; A0388B4F4A3652AE CRC64;

Query Match 100.0%; Score 49; DB 5; Length 90;

Best Local Similarity 23.5%; Pred. No. 1.2e+02;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17

Db 63 PKDDVCCFTVECLKRCF 79

RESULT 39

Q9BPJ4
ID Q9BPJ4 PRELIMINARY; PRT; 90 AA.
AC Q9BPJ4;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE CONOTOXIN SCAFFOLD III/IV.
OS Conus textile (Cloth-of-gold cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6494;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21105969; PubMed=11158371;
RA Canticello S.G., Gilad Y., Avdian N., Ben-Asher E., Levy Z.,
FA Inzilibar M.;
RT "Mechanisms for Evolving Hypervariability: The Case of Conopeptides.";
RL Mol. Biol. Evol. 18:120-131(2001).
DR EMBL; AF214929; AAG60357.1; -.
SQ SEQUENCE 90 AA; 10030 MW; 08EAF35A0CAAE36B CRC64;

Query Match 100.0%; Score 49; DB 5; Length 90;

Best Local Similarity 23.5%; Pred. No. 1.2e+02;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17

#####|:|:|:

Db 70 SRGRCGPTACMAGCR 86

RESULT 40

Q9BYR4
ID Q9BYR4 PRELIMINARY; PRT; 98 AA.
AC Q9BYR4;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)
DE KERATIN ASSOCIATED PROTEIN 4.3 (FRAGMENT).
GN KRTAP4.3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SCALP;
RA Rogers M.A., Langbein L., Winter H., Ehmman C., Korn B., Schweizer J.;
RT "Characterization of a cluster of human high/ ultrahigh keratin
associated proteins on chromosome 17q12-21.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ406935; CAC27574.1; -.
DR InterPro; IPR002494; Keratin_B2.
DR InterPro; IPR001368; TnFR_c6.
DR Pfam; PF01500; Keratin_B2; 1.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 98 AA; 10395 MW; 9F20C7A4113B1770 CRC64;

Query Match 100.0%; Score 49; DB 4; Length 98;

Best Local Similarity 23.5%; Pred. No. 1.3e+02;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17

#####|:|:|:

Db 2 CCRPSCCISCCRPSCC 18

RESULT 41

Q9ERH9
ID Q9ERH9 PRELIMINARY; PRT; 101 AA.

Q9ERH9;
AC 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE CYSTEINE-STRING PROTEIN (FRAGMENT).
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ORGAN OF CORTI;
RA Eybalin M., Renard N., Aure F., Safieddine S.;
RT "Cysteine-string protein, localized to vesicles associated to the
synaptic ribbon in inner hair cells, may be required for the onset of
hearing.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF305092; AAG16752.1; -.
FT NON_TER 1
FT NON_TER 101
SQ SEQUENCE 101 AA; 11188 MW; C37C13F7534FD823 CRC64;

Query Match 100.0%; Score 49; DB 11; Length 101;

Best Local Similarity 23.5%; Pred. No. 1.3e+02;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17

#####|:|:|:

Db 35 CGLLTCCYCCCGCCG 51

RESULT 42

Q9BYP8
ID Q9BYP8 PRELIMINARY; PRT; 105 AA.
AC Q9BYP8;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE KERATIN ASSOCIATED PROTEIN 16.1.
GN KRTAP16.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SCALP;
RA Rogers M.A., Langbein L., Winter H., Ehmman C., Korn B., Schweizer J.;
RT "Characterization of a cluster of human high/ ultrahigh keratin
associated proteins on chromosome 17q12-21.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ406952; CAC27591.1; -.
DR HSSP; P02876; 9WGA.
DR InterPro; IPR001007; VWFC.
DR PROSITE; PS01208; VWFC; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 105 AA; 9504 MW; AA633F91BD800582 CRC64;

Query Match 100.0%; Score 49; DB 4; Length 105;

Best Local Similarity 23.5%; Pred. No. 1.4e+02;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17

#####|:|:|:

Db 13 TOEQNCCECCCGGCC 29

RESULT 43

Q9BYU7
ID Q9BYU7 PRELIMINARY; PRT; 106 AA.
AC Q9BYU7;
DT 01-JUN-2001 (TReMBLrel. 17, Created)

```
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE KERATIN ASSOCIATED PROTEIN.
GN KRTAP4.13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SCALP;
RA Rogers M.A., Langbein L., Winter H., Ehmann C., Korn B., Schweizer J.;
RT "Characterization of a cluster of human high/ultrahigh keratin
associated proteins on chromosome 17q12-21."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ296168; CAC27563.1; -
DR InterPro; IPR002494; Keratin_B2.
DR Pfam; PF01500; Keratin_B2; 1.
DR PROSITE; PS00652; TNFR_NGFR.1; UNKNOWN.1.
SQ SEQUENCE 106 AA; 11570 MW; 15CF43A4C80C4375 CRC64;

Query Match 100.0%; Score 49; DB 4; Length 106;
Best Local Similarity 23.5%; Pred. No. 1.4e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 15 CGLENCRPSYCQTCC 31

RESULT 44
Q96RJ8 ID Q96RJ8 PRELIMINARY; PRT; 109 AA.
AC Q96RJ8;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE GAMMA CYSTEINE STRING PROTEIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tobaben S., Stahl B.;
RT "Novel isoforms of the cysteine string protein."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF368277; AAK60572.1; -
FT NON_TER 1 1
FT NON_TER 109 109
SQ SEQUENCE 109 AA; 12412 MW; 756384BC0FF089E3 CRC64;

Query Match 100.0%; Score 49; DB 4; Length 109;
Best Local Similarity 23.5%; Pred. No. 1.4e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 75 CTLLTCCFCFCCCCFC 91

RESULT 45
Q9D227 ID Q9D227 PRELIMINARY; PRT; 109 AA.
AC Q9D227;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE A030006P16RIK PROTEIN.
GN A030006P16RIK.
```

```
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=C57BL/6J; TISSUE=SKIN;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Hoffell D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK020698; BAB32179.1; -
DR HSSP; P02876; 9WGA.
DR MGD; MGI:1925164; A030006P16RIK.
DR InterPro; IPR000561; EGF-like.
DR PROSITE; PS00022; EGF_1; UNKNOWN.1.
SQ SEQUENCE 109 AA; 9655 MW; A5067E9BA0D2AE3C CRC64;
```

Query Match 100.0%; Score 49; DB 11; Length 109;
Best Local Similarity 23.5%; Pred. No. 1.4e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

QY 1 XXXXXCCXXXCCXXXCX 17
DB 13 SQEDCCCECCCCQGCC 29

```
RESULT 46
Q9BYQ1 ID Q9BYQ1 PRELIMINARY; PRT; 111 AA.
AC Q9BYQ1;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE KERATIN ASSOCIATED PROTEIN 9.5 (FRAGMENT).
GN KRTAP9.5.
```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE FROM N.A.
RC TISSUE=SCALP;
RA Rogers M.A., Langbein L., Winter H., Ehmann C., Korn B., Schweizer J.;
RT "Characterization of a cluster of human high/ultrahigh keratin
associated proteins on chromosome 17q12-21."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ406949; CAC27588.1; -
DR InterPro; IPR002494; Keratin_B2.
DR Pfam; PF01500; Keratin_B2; 1.
FT NON_TER 1 1
SQ SEQUENCE 111 AA; 11488 MW; 339AA429CB64CA7F CRC64;

Query Match 100.0%; Score 49; DB 4; Length 111;

Best Local Similarity 23.5%; Pred. No. 1.4e+02;
Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCCXXXCCXXXCX 17

Db 7 TCLTSCCQPCSCSTWCC 23

RESULT 47

OL14634 ID O14634 PRELIMINARY; PRT; 112 AA.

AC O14634; 01-JAN-1998 (TrEMBLrel. 05, Created)

DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

DE SKIN-SPECIFIC PROTEIN (FRAGMENT).

GN XP12.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

[1]

SEQUENCE FROM N.A.

RC TISSUE=SKIN;

RX MEDLINE=98008911; PubMed=9344646;

RA Zhao X.P., Elder J.T.;

RT "Positional cloning of novel skin-specific genes from the human

epidermal differentiation complex.";

RL Genomics 45:250-258(1997).

DR EMBL; AF005081; AAB83961.1; -

DR InterPro; IPR000561; EGF-like.

DR PROSITE; PS00022; EGF_1; UNKNOWN_1.

FT NON_TER 1

SQ SEQUENCE 112 AA; 11329 MW; 96826DC3E7362E6B CRC64;

Query Match 100.0%; Score 49; DB 4; Length 112;

Best Local Similarity 23.5%; Pred. No. 1.4e+02;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCCXXXCCXXXCX 17

Db 74 SGSGCCGCGGCGSCG 90

RESULT 48

Q9D228 ID Q9D228 PRELIMINARY; PRT; 113 AA.

AC Q9D228;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE A030005L19RIK PROTEIN.

GN A030005L19RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

[1]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=SKIN;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

Alizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,

Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavani T.,

Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

Schriml L.M., Staib F., Suzuki R., Tomita M., Wagner L., Washio T.,

Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

DR EMBL; AK020696; BAB32178.1; -

DR HSSP; P02876; 9WGA.

DR MGD; MGI:1925172; A030005L19RIK.

DR InterPro; IPR000561; EGF-like.

DR PROSITE; PS00022; EGF_1; UNKNOWN_1.

SQ SEQUENCE 113 AA; 10002 MW; A59A17886401FDA7 CRC64;

Query Match 100.0%; Score 49; DB 11; Length 113;

Best Local Similarity 23.5%; Pred. No. 1.4e+02;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCCXXXCCXXXCX 17

Db 68 KPTVVCCRRSCRCSCG 84

RESULT 49

Q9BYQ9 ID Q9BYQ9 PRELIMINARY; PRT; 114 AA.

AC Q9BYQ9;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)

DE KERATIN ASSOCIATED PROTEIN 4.8 (FRAGMENT).

GN KRTAP4.8.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

[1]

SEQUENCE FROM N.A.

RC TISSUE=SCALP;

RA Rogers M.A., Langbein L., Winter H., Ehmann C., Korn B., Schweizer J.;

RT "Characterization of a cluster of human high/ ultrahigh keratin

associated proteins on chromosome 17q12-21.";

RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ406940; CAC27579.1; -

DR InterPro; IPR002494; Keratin_B2.

DR InterPro; IPR001368; TNFR_c6.

DR Pfam; PF01500; Keratin_B2; 1.

DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.

DR PROSITE; PS01208; VWFC; UNKNOWN_1.

FT NON_TER 1

SQ SEQUENCE 114 AA; 12027 MW; C7E55BCC4690C245 CRC64;

Query Match 100.0%; Score 49; DB 4; Length 114;

Best Local Similarity 23.5%; Pred. No. 1.5e+02;

Matches 4; Conservative 13; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXCCXXXCCXXXCX 17

Db 3 CCKPOCCQSVCCQPTCC 19

RESULT 50

Q9RNT5 ID Q9RNT5 PRELIMINARY; PRT; 120 AA.

AC Q9RNT5;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 12:24:52 ; Search time 51.6 Seconds
(without alignments)
75.341 Million cell updates/sec

Title: US-09-493-795A-236

Perfect score: 197

Sequence: 1 FDGRNAPADDKASDLIAQIVRRACCSDDRRRCRCG 35

Scoring table:

BLOSUM62DX

Gapop 10.0 , Gapext 0.5

arched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 197 | 100.0 | 35 | 21 | AA21554 |
| 2 | 131 | 66.5 | 35 | 21 | AA21559 |
| 3 | 120 | 60.9 | 39 | 21 | AA21628 |
| 4 | 118 | 59.9 | 39 | 21 | AA21630 |
| 5 | 115 | 58.4 | 39 | 21 | AA21599 |
| 6 | 115 | 58.4 | 39 | 21 | AA21629 |
| 7 | 106 | 53.8 | 39 | 21 | AA21627 |
| 8 | 98.5 | 50.0 | 41 | 21 | AA21570 |
| 9 | 93.5 | 47.5 | 63 | 21 | AA21439 |
| 10 | 92 | 46.7 | 39 | 21 | AA21626 |
| 11 | 90.5 | 45.9 | 41 | 21 | AA21595 |

| | | | | | | |
|----|------|------|----|----|----------|--------------------|
| 12 | 90.5 | 45.9 | 41 | 21 | AA21598 | Cone snail alpha-c |
| 13 | 90 | 45.7 | 41 | 21 | AA21574 | Cone snail alpha-c |
| 14 | 90 | 45.7 | 41 | 21 | AA21575 | Cone snail alpha-c |
| 15 | 90 | 45.7 | 64 | 21 | AA21452 | Cone snail alpha-c |
| 16 | 88.5 | 44.9 | 42 | 21 | AA21603 | Cone snail alpha-c |
| 17 | 87.5 | 44.4 | 41 | 21 | AA21633 | Cone snail alpha-c |
| 18 | 87.5 | 44.4 | 62 | 21 | AA215159 | Alpha-conotoxin pr |
| 19 | 87 | 44.2 | 35 | 21 | AA21560 | Cone snail alpha-c |
| 20 | 86 | 43.7 | 61 | 21 | AA21450 | Cone snail alpha-c |
| 21 | 85.5 | 43.4 | 41 | 21 | AA21608 | Cone snail alpha-c |
| 22 | 84.5 | 42.9 | 38 | 21 | AA21609 | Cone snail alpha-c |
| 23 | 84.5 | 42.9 | 60 | 21 | AA21435 | Cone snail alpha-c |
| 24 | 84.5 | 42.9 | 65 | 18 | AA24901 | C. bandanus alpha- |
| 25 | 84 | 42.6 | 39 | 21 | AA21601 | Cone snail alpha-c |
| 26 | 83.5 | 42.4 | 38 | 21 | AA21606 | Cone snail alpha-c |
| 27 | 83.5 | 42.4 | 55 | 21 | AA21572 | Cone snail alpha-c |
| 28 | 83 | 42.1 | 35 | 21 | AA21555 | Cone snail alpha-c |
| 29 | 82.5 | 41.9 | 62 | 21 | AA21565 | Alpha-conotoxin pr |
| 30 | 82 | 41.6 | 35 | 21 | AA21553 | Cone snail alpha-c |
| 31 | 81.5 | 41.4 | 40 | 21 | AA21623 | Cone snail alpha-c |
| 32 | 81.5 | 41.4 | 41 | 21 | AA21611 | Cone snail alpha-c |
| 33 | 81.5 | 41.4 | 60 | 21 | AA21470 | Cone snail alpha-c |
| 34 | 81 | 41.1 | 32 | 21 | AA21558 | Cone snail alpha-c |
| 35 | 81 | 41.1 | 35 | 21 | AA21556 | Cone snail alpha-c |
| 36 | 81 | 41.1 | 39 | 21 | AA21590 | Cone snail alpha-c |
| 37 | 81 | 41.1 | 62 | 21 | AA21441 | Cone snail alpha-c |
| 38 | 80.5 | 40.9 | 41 | 21 | AA21579 | Cone snail alpha-c |
| 39 | 80.5 | 40.9 | 62 | 21 | AA21518 | Alpha-conotoxin pr |
| 40 | 80.5 | 40.9 | 63 | 21 | AA21426 | Cone snail alpha-c |
| 41 | 80.5 | 40.9 | 63 | 21 | AA21448 | Cone snail alpha-c |
| 42 | 80.5 | 40.9 | 63 | 21 | AA21473 | Cone snail alpha-c |
| 43 | 79.5 | 40.4 | 36 | 21 | AA21596 | Cone snail alpha-c |
| 44 | 78 | 39.6 | 35 | 21 | AA21557 | Cone snail alpha-c |
| 45 | 78 | 39.6 | 39 | 21 | AA21636 | Cone snail alpha-c |

ALIGNMENTS

RESULT 1

AA21554
ID AAB21554 standard; Peptide: 35 AA.
XX
AC AAB21554;
XX
DT 19-JAN-2001 (first entry)
XX
DE Cone snail alpha-conotoxin SEQ ID NO: 236.
XX
KW Cone snail; alpha-conotoxin; venom: disulphide bond; mood disorder;
KW neuronal nicotinic acetylcholine receptor; cardiovascular disorder;
KW gastric motility disorder; urinary incontinence; nicotine addiction;
KW small cell lung carcinoma.
XX
OS Conus imperialis.
XX
PN WO200044776-A1.
XX
PD 03-AUG-2000.
XX
PF 28-JAN-2000; 2000WO-US01979.
XX
PR 29-JAN-1999; 99US-0118381.
XX
PA (UTAH) UNIV UTAH RES FOUND.
XX
PA (COGN-) COGNEX INC.
XX
PI Watkins M, Olivera BM, Hillyard DR, McIntosh JM, Jones RM;
XX
XX WPI; 2000-505965/45.
DR N-PSDB; AAA89450.
XX
PT alpha-conotoxin polypeptides derived from the venom of cone snails

Qy 2 DGRNAPADDKASDLIAQIVRRACCSDRRCWR---CG 35

AA
AC
XX

DT 19-JAN-2001 (first entry)
XX Cone snail alpha-conotoxin SEQ ID NO: 386.
DE XX
XX Cone snail; alpha-conotoxin; venom; disulphide bond; mood disorder;
KW neuronal nicotinic acetylcholine receptor; cardiovascular disorder;
KW gastric motility disorder; urinary incontinence; nicotine addiction;
KW small cell lung carcinoma.
XX
XX Conus regius.
XX WO200044776-A1.
PN 03-AUG-2000.
PD 28-JAN-2000; 2000WO-US01979.
PE 29-JAN-1999; 99US-0118381.
XX (UTAH) UNIV UTAH RES FOUND.
XX (COGN-) COGNETIX INC.
PA Watkins M, Olivera BM, Hillyard DR, McIntosh JM, Jones RM;
PI WPI; 2000-505965/45.
XX N-PSDB; AAA89525.
DR alpha-conotoxin polypeptides derived from the venom of cone snails
XX useful e.g. as neuromuscular blocking agents for use in surgery and for
PT treating unipolar depression -
PT
XX Claim 39; Page 64; 229pp; English.
PS The present invention relates to a number of alpha-conotoxin peptides and
XX their coding sequences from a number of different species of cone snail.
XX These peptides are found in minute quantities in the venom of the snails,
CC and are targeted at the neuronal nicotinic acetylcholine receptors of the
CC nervous system. They usually contain two disulphide bonds, which give
CC them defined conformations, a rarity in molecules this small. The
CC alpha-conotoxins can be used as neuromuscular blocking agents in surgery,
CC and for treating disorders regulated at the neuronal nicotinic
CC acetylcholine receptors, including cardiovascular disorders, gastric
CC motility disorders, urinary incontinence, nicotine addiction, mood
CC disorders such as bipolar disorder, unipolar depression, dysthymia and
CC seasonal affective disorder, and small cell lung carcinoma.
XX Sequence 39 AA;
SQ

Query Match 58.4%; Score 115; DB 21; Length 39;
Best Local Similarity 75.9%; Pred. No. 5.5e-09;
Matches 22; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 2 DGRNAPADKASDLIAQIVRRACCSRRRC 30
DB 2 dgrnaadnrasdliaqivrgcshpac 30
||||| ||::||| ||||| ||| |
|

RESULT 7
AAB21627
ID AAB21627 standard; Peptide; 39 AA.
XX
XX AAB21627;
AC
XX 19-JAN-2001 (first entry)
DT
XX Cone snail alpha-conotoxin SEQ ID NO: 382.
DE
XX Cone snail; alpha-conotoxin; venom; disulphide bond; mood disorder;
KW neuronal nicotinic acetylcholine receptor; cardiovascular disorder;
KW gastric motility disorder; urinary incontinence; nicotine addiction;
KW small cell lung carcinoma.
XX

OS Conus regius.
XX WO200044776-A1.
PN 03-AUG-2000.
PD 28-JAN-2000; 2000WO-US01979.
XX 29-JAN-1999; 99US-0118381.
XX (UTAH) UNIV UTAH RES FOUND.
XX (COGN-) COGNETIX INC.
PA Watkins M, Olivera BM, Hillyard DR, McIntosh JM, Jones RM;
PI WPI; 2000-505965/45.
XX N-PSDB; AAA89523.
DR alpha-conotoxin polypeptides derived from the venom of cone snails
XX useful e.g. as neuromuscular blocking agents for use in surgery and for
PT treating unipolar depression -
PT
XX Claim 39; Page 64; 229pp; English.
PS The present invention relates to a number of alpha-conotoxin peptides and
XX their coding sequences from a number of different species of cone snail.
XX These peptides are found in minute quantities in the venom of the snails,
CC and are targeted at the neuronal nicotinic acetylcholine receptors of the
CC nervous system. They usually contain two disulphide bonds, which give
CC them defined conformations, a rarity in molecules this small. The
CC alpha-conotoxins can be used as neuromuscular blocking agents in surgery,
CC and for treating disorders regulated at the neuronal nicotinic
CC acetylcholine receptors, including cardiovascular disorders, gastric
CC motility disorders, urinary incontinence, nicotine addiction, mood
CC disorders such as bipolar disorder, unipolar depression, dysthymia and
CC seasonal affective disorder, and small cell lung carcinoma.
XX Sequence 39 AA;
SQ

Query Match 53.8%; Score 106; DB 21; Length 39;
Best Local Similarity 66.7%; Pred. No. 9.9e-08;
Matches 20; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
QY 2 DGRNAPADKASDLIAQIVRRACCSRRRC 31
DB 2 dgrntaadekasdlisqtkvrdccshplcr 31
||||| ||::||| ||||| ||| |
|

RESULT 8
AAB21570
ID AAB21570 standard; Peptide; 41 AA.
XX
XX AAB21570;
AC
XX 19-JAN-2001 (first entry)
DT
XX Cone snail alpha-conotoxin SEQ ID NO: 268.
DE
XX Cone snail; alpha-conotoxin; venom; disulphide bond; mood disorder;
KW neuronal nicotinic acetylcholine receptor; cardiovascular disorder;
KW gastric motility disorder; urinary incontinence; nicotine addiction;
KW small cell lung carcinoma.
XX
XX Conus betulinus.
XX WO200044776-A1.
PN 03-AUG-2000.
PD 28-JAN-2000; 2000WO-US01979.
XX 29-JAN-1999; 99US-0118381.
XX

XX (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
XX Watkins M, Olivera BM, Hillyard DR, McIntosh JM, Jones RM;
XX
XX WPI; 2000-505965/45.
DR N-PSDB; AAA89466.
XX
XX alpha-conotoxin polypeptides derived from the venom of cone snails
PT useful e.g. as neuromuscular blocking agents for use in surgery and for
PT treating unipolar depression -
XX
XX Claim 39; Page 50; 229pp; English.
PS
XX The present invention relates to a number of alpha-conotoxin peptides and
CC their coding sequences from a number of different species of cone snail.
CC These peptides are found in minute quantities in the venom of the snails,
CC and are targeted at the neuronal nicotinic acetylcholine receptors of the
CC nervous system. They usually contain two disulphide bonds, which give
CC them defined conformations, a rarity in molecules this small. The
CC alpha-conotoxins can be used as neuromuscular blocking agents in surgery,
CC and for treating disorders regulated at the neuronal nicotinic
CC acetylcholine receptors, including cardiovascular disorders, gastric
CC motility disorders, urinary incontinence, nicotine addiction, mood
CC disorders such as bipolar disorder, unipolar depression, dysthymia and
CC seasonal affective disorder, and small cell lung carcinoma.
XX Sequence 41 AA;
SQ

Query Match 50.0%; Score 98.5; DB 21; Length 41;
Best Local Similarity 52.6%; Pred. No. 1.2e-06;
Matches 20; Conservative 4; Mismatches 11; Indels 3; Gaps 1;
Qy 1 FDGRNAPADKASDLIAQIVRRACCSDRRC---RWRCG 35
Db 1 ydgrnaaaddkafdlamtirggccsyppciasnpkcg 38
:||||| ||||| ||| :| ||| : |||
:||||| ||||| ||| :| ||| : |||

RESULT 9
AAB21439
ID AAB21439 standard; Protein; 63 AA.
AC AAB21439;
XX
XX 19-JAN-2001 (first entry)
DE Cone snail alpha-conotoxin SEQ ID NO: 85.
XX
XX Cone snail; alpha-conotoxin; venom; disulphide bond; mood disorder;
KW neuronal nicotinic acetylcholine receptor; cardiovascular disorder;
KW gastric motility disorder; urinary incontinence; nicotine addiction;
KW small cell lung carcinoma.
XX
XX Conus characteristic.
XX
XX WO200044776-A1.
PN
XX
XX 03-AUG-2000.
PD
XX
XX 28-JAN-2000; 2000WO-US01979.
PF
XX
XX 29-JAN-1999; 99US-0118381.
PR
XX
XX (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
XX Watkins M, Olivera BM, Hillyard DR, McIntosh JM, Jones RM;
PI
XX
XX WPI; 2000-505965/45.
DR N-PSDB; AAA89414.
XX

PT alpha-conotoxin polypeptides derived from the venom of cone snails
PT useful e.g. as neuromuscular blocking agents for use in surgery and for
PT treating unipolar depression -
XX
XX Claim 39; Page 35; 229pp; English.
PS
XX The present invention relates to a number of alpha-conotoxin peptides and
CC their coding sequences from a number of different species of cone snail.
CC These peptides are found in minute quantities in the venom of the snails,
CC and are targeted at the neuronal nicotinic acetylcholine receptors of the
CC nervous system. They usually contain two disulphide bonds, which give
CC them defined conformations, a rarity in molecules this small. The
CC alpha-conotoxins can be used as neuromuscular blocking agents in surgery,
CC and for treating disorders regulated at the neuronal nicotinic
CC acetylcholine receptors, including cardiovascular disorders, gastric
CC motility disorders, urinary incontinence, nicotine addiction, mood
CC disorders such as bipolar disorder, unipolar depression, dysthymia and
CC seasonal affective disorder, and small cell lung carcinoma.
XX Sequence 63 AA;
SQ
Query Match 47.5%; Score 93.5; DB 21; Length 63;
Best Local Similarity 55.3%; Pred. No. 9.1e-06;
Matches 21; Conservative 4; Mismatches 8; Indels 5; Gaps 2;
Qy 2 DGRNAPADKASDLIAQIVRRACCSDRRCRWR---CG 35
Db 24 egrnaaakkasdlvaltv-rgccaicreqlnaaycg 60
:||||| ||||| :| ||| : ||| : |||
:||||| ||||| :| ||| : ||| : |||

RESULT 10
AAB21626
ID AAB21626 standard; Peptide; 39 AA.
XX
XX AAB21626;
XX
XX 19-JAN-2001 (first entry)
XX
XX Cone snail alpha-conotoxin SEQ ID NO: 380.
DE
XX
XX Cone snail; alpha-conotoxin; venom; disulphide bond; mood disorder;
KW neuronal nicotinic acetylcholine receptor; cardiovascular disorder;
KW gastric motility disorder; urinary incontinence; nicotine addiction;
KW small cell lung carcinoma.
XX
XX Conus regius.
OS
XX
XX WO200044776-A1.
PN
XX
XX 03-AUG-2000.
PD
XX
XX 28-JAN-2000; 2000WO-US01979.
PF
XX
XX 29-JAN-1999; 99US-0118381.
PR
XX
XX (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.
XX
XX Watkins M, Olivera BM, Hillyard DR, McIntosh JM, Jones RM;
PI
XX
XX WPI; 2000-505965/45.
DR N-PSDB; AAA89522.
XX
XX alpha-conotoxin polypeptides derived from the venom of cone snails
PT useful e.g. as neuromuscular blocking agents for use in surgery and for
PT treating unipolar depression -
XX
XX Claim 39; Page 63; 229pp; English.
PS
XX
XX The present invention relates to a number of alpha-conotoxin peptides and
CC their coding sequences from a number of different species of cone snail.
CC These peptides are found in minute quantities in the venom of the snails,
CC and are targeted at the neuronal nicotinic acetylcholine receptors of the
CC nervous system. They usually contain two disulphide bonds, which give
CC them defined conformations, a rarity in molecules this small. The
CC alpha-conotoxins can be used as neuromuscular blocking agents in surgery,
CC and for treating disorders regulated at the neuronal nicotinic
CC acetylcholine receptors, including cardiovascular disorders, gastric
CC motility disorders, urinary incontinence, nicotine addiction, mood
CC disorders such as bipolar disorder, unipolar depression, dysthymia and
CC seasonal affective disorder, and small cell lung carcinoma.
XX Sequence 63 AA;
SQ

The present invention relates to a number of alpha-conotoxin peptides and their coding sequences from a number of different species of cone snail. These peptides are found in minute quantities in the venom of the snails, and are targeted at the neuronal nicotinic acetylcholine receptors of the nervous system. They usually contain two disulphide bonds, which give them defined conformations, a rarity in molecules this small. The alpha-conotoxins can be used as neuromuscular blocking agents in surgery, and for treating disorders regulated at the neuronal nicotinic acetylcholine receptors, including cardiovascular disorders, gastric motility disorders, urinary incontinence, nicotine addiction, mood disorders such as bipolar disorder, unipolar depression, dysthymia and seasonal affective disorder, and small cell lung carcinoma.

Db 2 dgrnaandgasdlaatv-rgccavpsclrnpldgcg 38
 ||||| |:|:|:|:|:| | | | | | | | | |

RESULT 13

AAB21574
 ID AAB21574 standard; Peptide: 41 AA.

XX AAB21574;

DT 19-JAN-2001 (first entry)

XX Cone snail alpha-conotoxin SEQ ID NO: 276.

XX Cone snail; alpha-conotoxin; venom; disulphide bond; mood disorder;
 KW neuronal nicotinic acetylcholine receptor; cardiovascular disorder;
 KW gastric motility disorder; urinary incontinence; nicotine addiction;
 KW small cell lung carcinoma.

XX Conus radiatus.

XX WO200044776-A1.

XX 03-AUG-2000.

XX 28-JAN-2000; 2000WO-US01979.

XX 29-JAN-1999; 99US-0118381.

XX (UTAH) UNIV UTAH RES FOUND.

XX (COGN-) COGNETIX INC.

XX Watkins M, Olivera BM, Hillyard DR, McIntosh JM, Jones RM;

XX WPI; 2000-505965/45.

XX N-PSDB; AAA89470.

XX alpha-conotoxin polypeptides derived from the venom of cone snails
 PT useful e.g. as neuromuscular blocking agents for use in surgery and for
 PT treating unipolar depression -

XX Claim 39; Page 51; 229pp; English.

XX The present invention relates to a number of alpha-conotoxin peptides and
 CC their coding sequences from a number of different species of cone snail.
 CC These peptides are found in minute quantities in the venom of the snails,
 CC and are targeted at the neuronal nicotinic acetylcholine receptors of the
 CC nervous system. They usually contain two disulphide bonds, which give
 CC them defined conformations, a rarity in molecules this small. The
 CC alpha-conotoxins can be used as neuromuscular blocking agents in surgery,
 CC and for treating disorders regulated at the neuronal nicotinic
 CC acetylcholine receptors, including cardiovascular disorders, gastric
 CC motility disorders, urinary incontinence, nicotine addiction, mood
 CC disorders such as bipolar disorder, unipolar depression, dysthymia and
 CC seasonal affective disorder, and small cell lung carcinoma.

XX Sequence 41 AA;

Query Match

Best Local Similarity 45.7%; Score 90; DB 21; Length 41;
 Matches 18; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 FDGRNAPADDKASDLIAQIVRRACCSDRRC 30

Db 1 fdgrnaaadykgseallamtvrvgccsyppc 30

RESULT 14

AAB21575

ID AAB21575 standard; Peptide: 41 AA.

XX AAB21575;

XX

DT 19-JAN-2001 (first entry)

XX Cone snail alpha-conotoxin SEQ ID NO: 278.

XX Cone snail; alpha-conotoxin; venom; disulphide bond; mood disorder;
 KW neuronal nicotinic acetylcholine receptor; cardiovascular disorder;
 KW gastric motility disorder; urinary incontinence; nicotine addiction;
 KW small cell lung carcinoma.

XX Conus radiatus.

XX WO200044776-A1.

XX 03-AUG-2000.

XX 28-JAN-2000; 2000WO-US01979.

XX 29-JAN-1999; 99US-0118381.

XX (UTAH) UNIV UTAH RES FOUND.

XX (COGN-) COGNETIX INC.

XX Watkins M, Olivera BM, Hillyard DR, McIntosh JM, Jones RM;

XX WPI; 2000-505965/45.

XX N-PSDB; AAA89471.

XX alpha-conotoxin polypeptides derived from the venom of cone snails
 PT useful e.g. as neuromuscular blocking agents for use in surgery and for
 PT treating unipolar depression -

XX Claim 39; Page 51; 229pp; English.

XX The present invention relates to a number of alpha-conotoxin peptides and
 CC their coding sequences from a number of different species of cone snail.
 CC These peptides are found in minute quantities in the venom of the snails,
 CC and are targeted at the neuronal nicotinic acetylcholine receptors of the
 CC nervous system. They usually contain two disulphide bonds, which give
 CC them defined conformations, a rarity in molecules this small. The
 CC alpha-conotoxins can be used as neuromuscular blocking agents in surgery,
 CC and for treating disorders regulated at the neuronal nicotinic
 CC acetylcholine receptors, including cardiovascular disorders, gastric
 CC motility disorders, urinary incontinence, nicotine addiction, mood
 CC disorders such as bipolar disorder, unipolar depression, dysthymia and
 CC seasonal affective disorder, and small cell lung carcinoma.

XX Sequence 41 AA;

Query Match

Best Local Similarity 45.7%; Score 90; DB 21; Length 41;
 Matches 18; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 FDGRNAPADDKASDLIAQIVRRACCSDRRC 30

Db 1 fdgrnaaadykgseallamtvrvgccsyppc 30

RESULT 15

AAB21452

ID AAB21452 standard; Protein: 64 AA.

XX AAB21452;

DT 19-JAN-2001 (first entry)

XX Cone snail alpha-conotoxin SEQ ID NO: 111.

XX Cone snail; alpha-conotoxin; venom; disulphide bond; mood disorder;
 KW neuronal nicotinic acetylcholine receptor; cardiovascular disorder;
 KW gastric motility disorder; urinary incontinence; nicotine addiction;
 KW small cell lung carcinoma.

XX OS Conus betulinus.
XX PN WO200044776-A1.
XX PD 03-AUG-2000.
XX PF 28-JAN-2000; 2000WO-US01979.
XX PR 29-JAN-1999; 99US-0118381.
XX PA (UTAH) UNIV UTAH RES FOUND.
XX PA (COGN-) COGNETIX INC.
XX PI Watkins M, Olivera BM, Hillyard DR, McIntosh JM, Jones RM;
XX DR WPI; 2000-505965/45.
XX DR N-PSDB; AAA98427.
XX PT alpha-conotoxin polypeptides derived from the venom of cone snails
XX PT useful e.g. as neuromuscular blocking agents for use in surgery and for
XX PT treating unipolar depression -
XX PS Claim 39; Page 39; 229pp; English.
XX CC The present invention relates to a number of alpha-conotoxin peptides and
XX CC their coding sequences from a number of different species of cone snail.
XX CC These peptides are found in minute quantities in the venom of the snails,
XX CC and are targeted at the neuronal nicotinic acetylcholine receptors of the
XX CC nervous system. They usually contain two disulphide bonds, which give
XX CC them defined conformations, a rarity in molecules this small. The
XX CC alpha-conotoxins can be used as neuromuscular blocking agents in surgery,
XX CC and for treating disorders regulated at the neuronal nicotinic
XX CC acetylcholine receptors, including cardiovascular disorders, gastric
XX CC motility disorders, urinary incontinence, nicotine addiction, mood
XX CC disorders such as bipolar disorder, unipolar depression, dysthymia and
XX CC seasonal affective disorder, and small cell lung carcinoma.
XX SQ Sequence 64 AA;

Query Match 45.7%; Score 90; DB 21; Length 64;
Best Local Similarity 60.0%; Pred. No. 2.9e-05;
Matches 18; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
Qy 1 FDGRNAPADDKASDLIAQIVRACCSDRRC 30
Db 23 frgrnsaandkrdsdlaalsvrgcshpac 52

Search completed: July 1, 2002, 12:24:52
Job time: 239 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 12:25:19 ; Search time 20.57 Seconds
(without alignments)
41.560 Million cell updates/sec

Title: US-09-493-795A-236
Perfect score: 197
Sequence: 1 FDGRNAPADKASDLIAQIVRRACCSDRRCWRG 35

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*

2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*

3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*

4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*

5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*

6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|---------------------|--------------------|
| 1 | 93.5 | 47.5 | 68 | 1 US-08-137-800-47 | Sequence 47, Appl |
| 2 | 93.5 | 47.5 | 68 | 1 US-08-477-383-47 | Sequence 47, Appl |
| 3 | 93.5 | 47.5 | 68 | 1 US-08-487-174-47 | Sequence 47, Appl |
| 4 | 93.5 | 47.5 | 68 | 1 US-08-480-750-47 | Sequence 47, Appl |
| 5 | 87.5 | 44.4 | 62 | 4 US-09-488-799-89 | Sequence 89, Appl |
| 6 | 84.5 | 42.9 | 65 | 1 US-08-137-800-46 | Sequence 46, Appl |
| 7 | 84.5 | 42.9 | 65 | 1 US-08-477-383-46 | Sequence 46, Appl |
| 8 | 84.5 | 42.9 | 65 | 1 US-08-487-174-46 | Sequence 46, Appl |
| 9 | 84.5 | 42.9 | 65 | 1 US-08-480-750-46 | Sequence 46, Appl |
| 10 | 82.5 | 41.9 | 62 | 4 US-09-488-799-101 | Sequence 101, Appl |
| 11 | 80.5 | 40.9 | 62 | 4 US-09-488-799-87 | Sequence 87, Appl |
| 12 | 80.5 | 40.9 | 68 | 1 US-08-137-800-48 | Sequence 48, Appl |
| 13 | 80.5 | 40.9 | 68 | 1 US-08-477-383-48 | Sequence 48, Appl |
| 14 | 80.5 | 40.9 | 68 | 1 US-08-487-174-48 | Sequence 48, Appl |
| 15 | 80.5 | 40.9 | 68 | 1 US-08-480-750-48 | Sequence 48, Appl |
| 16 | 75 | 38.1 | 62 | 1 US-08-137-800-50 | Sequence 50, Appl |
| 17 | 75 | 38.1 | 62 | 1 US-08-477-383-50 | Sequence 50, Appl |
| 18 | 75 | 38.1 | 62 | 1 US-08-487-174-50 | Sequence 50, Appl |
| 19 | 75 | 38.1 | 62 | 1 US-08-480-750-50 | Sequence 50, Appl |
| 20 | 74 | 37.6 | 62 | 1 US-08-137-800-51 | Sequence 51, Appl |
| 21 | 74 | 37.6 | 62 | 1 US-08-477-383-51 | Sequence 51, Appl |
| 22 | 74 | 37.6 | 62 | 1 US-08-487-174-51 | Sequence 51, Appl |
| 23 | 74 | 37.6 | 62 | 1 US-08-480-750-51 | Sequence 51, Appl |
| 24 | 70.5 | 35.8 | 63 | 4 US-09-488-799-93 | Sequence 93, Appl |
| 25 | 70.5 | 35.8 | 63 | 4 US-09-488-799-99 | Sequence 99, Appl |
| 26 | 65.5 | 33.2 | 63 | 4 US-09-488-799-97 | Sequence 97, Appl |
| 27 | 64 | 32.5 | 12 | 1 US-08-137-800-12 | Sequence 12, Appl |

Sequence 12, Appl
Sequence 12, Appl
Sequence 12, Appl
Sequence 13, Appl
Sequence 9, Appl
Sequence 91, Appl
Sequence 40, Appl
Sequence 40, Appl
Sequence 40, Appl
Sequence 40, Appl
Sequence 49, Appl
Sequence 49, Appl
Sequence 29, Appl
Sequence 13, Appl
Sequence 36, Appl
Sequence 45, Appl

ALIGNMENTS

RESULT 1
US-08-137-800-47
; Sequence 47, Application US/08137800
; Patent No. 5514774
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Hilliard, David R.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Santos, Ameurflina D.
; TITLE OF INVENTION: Conotoxin Peptides
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue N.W., Suite 1000
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/137,800
; FILING DATE: 19-OCT-1993
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24260-104763
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Conus caracteristicus
; US-08-137-800-47

Query Match 47.5%; Score 93.5; DB 1; Length 68;
Best Local Similarity 55.3%; Pred. No. 3.7e-06;
Matches 21; Conservative 4; Mismatches 8; Indels 5; Gaps 2;

STREET: 1201 New York Avenue, N.W., Suite 1000
CITY: Washington
STATE: DC
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,750
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/137,800
FILING DATE: 19-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/084,848
FILING DATE: 29-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24260-107673
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 68 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Conus characteristicus
US-08-480-750-47

Query Match 47.5%; Score 93.5; DB 1; Length 68;
Best Local Similarity 55.3%; Pred. No. 3.7e-06;
Matches 21; Conservative 4; Mismatches 8; Indels 5; Gaps 2;

QY 2 DGRNAPADKASDLIAQIVRRACCSDDRCRWR---CG 35
29 EGRNAAKDKASDLVLTV-RGCCAIRECLQNAAYCG 65

RESULT 5
US-09-488-799-89
Sequence 89, Application US/09488799
Patent No. 6268473
GENERAL INFORMATION:
APPLICANT: Olivera, Baldomero M.
APPLICANT: Layer, Richard T.
APPLICANT: Watkins, Maren
APPLICANT: Hillyard, David R.
APPLICANT: McIntosh, J. Michael
APPLICANT: Schoenfeld, Robert
APPLICANT: Jones, Robert M.
TITLE OF INVENTION: Alpha Conotoxin Peptides
FILE REFERENCE: Alphas 1
CURRENT APPLICATION NUMBER: US/09/488,799
CURRENT FILING DATE: 2000-01-21
EARLIER APPLICATION NUMBER: 60/116,881
EARLIER FILING DATE: 1999-01-22
EARLIER APPLICATION NUMBER: 60/116,882
EARLIER FILING DATE: 1999-01-22
NUMBER OF SEQ ID NOS: 101
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 89
LENGTH: 62

TYPE: PRT
ORGANISM: Conus ermineus
US-09-488-799-89

Query Match 44.4%; Score 87.5; DB 4; Length 62;
Best Local Similarity 63.3%; Pred. No. 2.2e-05;
Matches 19; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

QY 2 DGRNAPADKASDLIAQIVRR-ACCSDDRC 30
DB 24 DGRDAAANDKASDLIALTARRDPCCSNPAC 53

RESULT 6
US-08-137-800-46
Sequence 46, Application US/08137800
Patent No. 5514774
GENERAL INFORMATION:
APPLICANT: Olivera, Baldomero M.
APPLICANT: Cruz, Lourdes J.
APPLICANT: Hillyard, David R.
APPLICANT: McIntosh, J. Michael
APPLICANT: Santos, Ameurfin D.
TITLE OF INVENTION: Conotoxin Peptides
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti
STREET: 1201 New York Avenue N.W., Suite 1000
CITY: Washington
STATE: DC
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/137,800
FILING DATE: 19-OCT-1993
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24260-104763
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 65 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Conus bandanus
US-08-137-800-46

Query Match 42.9%; Score 84.5; DB 1; Length 65;
Best Local Similarity 62.1%; Pred. No. 6e-05;
Matches 18; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 2 DGRNAPADKASDLIAQIVRRACCSDDRC 30
DB 29 DGRNAAKDKASDLVLTV-KGCCSHAPAC 56

RESULT 7
US-08-477-383-46
Sequence 46, Application US/08477383

Patent No. 5589340
GENERAL INFORMATION:
APPLICANT: Olivera, Baldomero M.
APPLICANT: Cruz, Lourdes J.
APPLICANT: Hillyard, David R.
APPLICANT: Macintosh, J. Michael
APPLICANT: Santos, Aneurfino S.
TITLE OF INVENTION: Conotoxin Peptides
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti
STREET: 1201 New York Avenue, N.W., Suite 1000
CITY: Washington
STATE: DC
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,383
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/137,800
FILING DATE: 19-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/084,848
FILING DATE: 29-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24260-107673
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 65 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Conus bandanus
US-08-477-383-46

Query Match 42.98; Score 84.5; DB 1; Length 65;
Best Local Similarity 62.18; Pred. No. 6e-05;
Matches 18; Conservative 2; Mismatches 8; Indels 1; Gaps 1;
Qy 2 DGRNAPADKASDLIAQIVRRACSDRRRC 30
Db 29 DGRNAAKDRASDLVLTV-KGCCSHPAC 56

RESULT 8
US-08-487-174-46
Sequence 46, Application US/08487174
Patent No. 5595972
GENERAL INFORMATION:
APPLICANT: Olivera, Baldomero M.
APPLICANT: Cruz, Lourdes J.
APPLICANT: Hillyard, David R.
APPLICANT: Macintosh, J. Michael
APPLICANT: Santos, Aneurfino S.
TITLE OF INVENTION: Conotoxin Peptides
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESSES:

ADDRESSEE: Venable, Baetjer, Howard & Civiletti
STREET: 1201 New York Avenue, N.W., Suite 1000
CITY: Washington
STATE: DC
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,174
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/137,800
FILING DATE: 19-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/084,848
FILING DATE: 29-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24260-107673
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 65 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Conus bandanus
US-08-487-174-46

Query Match 42.98; Score 84.5; DB 1; Length 65;
Best Local Similarity 62.18; Pred. No. 6e-05;
Matches 18; Conservative 2; Mismatches 8; Indels 1; Gaps 1;
Qy 2 DGRNAPADKASDLIAQIVRRACSDRRRC 30
Db 29 DGRNAAKDRASDLVLTV-KGCCSHPAC 56

RESULT 9
US-08-480-750-46
Sequence 46, Application US/08480750
Patent No. 5633347
GENERAL INFORMATION:
APPLICANT: Olivera, Baldomero M.
APPLICANT: Cruz, Lourdes J.
APPLICANT: Hillyard, David R.
APPLICANT: Macintosh, J. Michael
APPLICANT: Santos, Aneurfino S.
TITLE OF INVENTION: Conotoxin Peptides
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti
STREET: 1201 New York Avenue, N.W., Suite 1000
CITY: Washington
STATE: DC
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

Query Match 41.9%; Score 82.5; DB 4; Length 62;
Best Local Similarity 60.0%; Pred. No. 0.00011;
Matches 18; Conservative 4; Mismatches 7; Indels

RESULT 12
US-08-137-800-48
; Sequence 48, Application US/08137800
; Patent No. 5514774
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Hilliard, David R.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Santos, Ameurina D.
; TITLE OF INVENTION: Conotoxin Peptides
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue N.W., Suite 1000
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/137,800
; FILING DATE: 19-OCT-1993
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957

REFERENCE/DOCKET NUMBER: 24260-104763
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 68 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Conus magus
US-08-137-800-48

Query Match 40.9%; Score 80.5; DB 1; Length 68;

Best Local Similarity 47.4%; Pred. No. 0.00023;
Matches 18; Conservative 5; Mismatches 10; Indels 5; Gaps 2;

2 DGRNAPADKASDLIAQIVRRACCSDRRCRWR---CG 35
||||| :||||| : : ||| : ||
29 DGRNAAANDKASDVIT-LALKGCCSNPVCHLEHSLNLCG 65

RESULT 13

US-08-477-383-48
Sequence 48, Application US/08477383
Patent No. 5589340
GENERAL INFORMATION:
APPLICANT: Olivera, Baldomero M.
APPLICANT: Cruz, Lourdes J.
APPLICANT: Hillyard, David R.
APPLICANT: Macintosh, J. Michael
APPLICANT: Santos, Ameurfino S.
TITLE OF INVENTION: Conotoxin Peptides
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti
STREET: 1201 New York Avenue, N.W., Suite 1000
CITY: Washington
STATE: DC
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,383
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER:
FILING DATE: 19-OCT-1993
APPLICATION NUMBER: US 08/137,800
FILING DATE: 19-OCT-1993
PRIOR APPLICATION NUMBER: US 08/084,848
FILING DATE: 29-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24260-107673
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 68 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear

MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Conus magus
US-08-477-383-48

Query Match 40.9%; Score 80.5; DB 1; Length 68;

Best Local Similarity 47.4%; Pred. No. 0.00023;
Matches 18; Conservative 5; Mismatches 10; Indels 5; Gaps 2;

QY 2 DGRNAPADKASDLIAQIVRRACCSDRRCRWR---CG 35
||||| :||||| : : ||| : ||
Db 29 DGRNAAANDKASDVIT-LALKGCCSNPVCHLEHSLNLCG 65

RESULT 14

US-08-487-174-48
Sequence 48, Application US/08487174
Patent No. 5595972
GENERAL INFORMATION:
APPLICANT: Olivera, Baldomero M.
APPLICANT: Cruz, Lourdes J.
APPLICANT: Hillyard, David R.
APPLICANT: Macintosh, J. Michael
APPLICANT: Santos, Ameurfino S.
TITLE OF INVENTION: Conotoxin Peptides
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti
STREET: 1201 New York Avenue, N.W., Suite 1000
CITY: Washington
STATE: DC
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,174
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION NUMBER:
FILING DATE: 19-OCT-1993
APPLICATION NUMBER: US 08/137,800
FILING DATE: 19-OCT-1993
PRIOR APPLICATION NUMBER: US 08/084,848
FILING DATE: 29-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24260-107673
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 68 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Conus magus
US-08-487-174-48

Query Match 40.9%; Score 80.5; DB 1; Length 68;

Best Local Similarity 47.4%; Pred. No. 0.00023;
Matches 18; Conservative 5; Mismatches 10; Indels 5; Gaps 2;

Qy 2 DGRNAPADKASDLIAQIVRRACCSDRRCWR-----CG 35
|||||1:|||||1: : |||: |
Db 29 DGRNAAANDKASDVIT-LALKGCCSNPVCHLEHSNLCG 65

RESULT 15
US-08-480-750-48
: Sequence 48, Application US/08480750
: Patent No. 5633347
: GENERAL INFORMATION:
: APPLICANT: Olivera, Baldomero M.
: APPLICANT: Cruz, Lourdes J.
: APPLICANT: Hillyard, David R.
: APPLICANT: Macintosh, J. Michael
: APPLICANT: Santos, Ameurfino S.
: TITLE OF INVENTION: Conotoxin Peptides
: NUMBER OF SEQUENCES: 59
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Venable, Baetjer, Howard & Civiletti
: STREET: 1201 New York Avenue, N.W., Suite 1000
: CITY: Washington
: STATE: DC
: COUNTRY: U.S.A.
: ZIP: 20005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/480.750
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/137.800
: FILING DATE: 19-OCT-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/084.848
: FILING DATE: 29-JUN-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Ihnen, Jeffrey L.
: REGISTRATION NUMBER: 28,957
: REFERENCE/DOCKET NUMBER: 24260-107673
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-962-4810
: TELEFAX: 202-962-8300
: INFORMATION FOR SEQ ID NO: 48:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 68 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: HYPOTHETICAL: NO
: ORIGINAL SOURCE:
: ORGANISM: Conus magus
US-08-480-750-48

Query Match 40.9%; Score 80.5; DB 1; Length 68;
Best Local Similarity 47.4%; Pred. No. 0.00023;
Matches 18; Conservative 5; Mismatches 10; Indels 5; Gaps 2;
Qy 2 DGRNAPADKASDLIAQIVRRACCSDRRCWR-----CG 35
|||||1:|||||1: : |||: |
Db 29 DGRNAAANDKASDVIT-LALKGCCSNPVCHLEHSNLCG 65

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 12:25:52 ; Search time 25.3 Seconds
(without alignments)
132.930 Million cell updates/sec

Title: US-09-493-795A-236

Perfect score: 197

Sequence: 1 FDGRNAPADDKASDLIAIVRRACCSDRRCRWRCG 35

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71.*

1: p1r1.*

2: p1r2.*

3: p1r3.*

4: p1r4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|----------|----------------------|
| 1 | 64 | 32.5 | 12 | 1 A53709 | alpha-conotoxin Im |
| 2 | 57 | 28.9 | 202 | 2 B45512 | cold-regulated prote |
| 3 | 57 | 28.9 | 521 | 2 T46250 | hypothetical prote |
| 4 | 54 | 27.4 | 470 | 2 AF2828 | FAD dependent oxid |
| 5 | 54 | 27.4 | 477 | 2 D97606 | glycolate oxidase |
| 6 | 52.5 | 26.6 | 72 | 2 A42325 | orf 5' to pheC - p |
| 7 | 52 | 26.4 | 252 | 2 A12372 | hypothetical prote |
| 8 | 52 | 26.4 | 565 | 2 AF3639 | tRNA pseudouridine |
| 9 | 52 | 26.4 | 611 | 2 AD2812 | dihydroxy-acid deh |
| 10 | 52 | 26.4 | 642 | 2 E97590 | dihydroxy-acid deh |
| 11 | 51.5 | 26.1 | 157 | 2 JN0057 | hypothetical 17K p |
| 12 | 51.5 | 26.1 | 534 | 2 C70842 | probable pmma prot |
| 13 | 51.5 | 26.1 | 665 | 2 T25228 | hypothetical prote |
| 14 | 51 | 25.9 | 360 | 2 D90657 | hypothetical prote |
| 15 | 51 | 25.9 | 360 | 2 D85008 | hypothetical prote |
| 16 | 51 | 25.9 | 863 | 2 D70770 | probable glycogen |
| 17 | 50.5 | 25.6 | 586 | 2 A85057 | probable transposo |
| 18 | 50.5 | 25.6 | 1552 | 2 T05408 | hypothetical prote |
| 19 | 50 | 25.4 | 174 | 2 B95940 | hypothetical prote |
| 20 | 50 | 25.4 | 414 | 2 B98354 | hypothetical prote |
| 21 | 50 | 25.4 | 869 | 2 AD2418 | SWI/SNF family hel |
| 22 | 49.5 | 25.1 | 176 | 2 H75332 | ankyrin-related pr |
| 23 | 49.5 | 25.1 | 230 | 2 H70357 | fumarate reductase |
| 24 | 49.5 | 25.1 | 418 | 2 B97252 | UDP-N-acetylglucos |
| 25 | 49.5 | 25.1 | 825 | 2 A26983 | regulatory protein |
| 26 | 49 | 24.9 | 88 | 2 S16161 | BLT14 protein - ba |
| 27 | 49 | 24.9 | 132 | 2 T02941 | CP12 protein precu |
| 28 | 49 | 24.9 | 145 | 2 I48093 | phospholipase A2 (|
| 29 | 49 | 24.9 | 860 | 2 F86349 | hypothetical prote |

30 49 24.9 975 2 T42576
31 48.5 24.6 957 2 AH2227
32 48 24.4 381 1 A47327
33 48 24.4 640 2 T08179
34 47.5 24.1 532 2 C87793
35 47.5 24.1 618 2 A87243
36 47.5 24.1 927 2 T00357
37 47 23.9 110 1 P3MLR5
38 47 23.9 128 2 T15101
39 47 23.9 129 2 S65803
40 47 23.9 189 2 C87606
41 47 23.9 261 2 T32399
42 47 23.9 306 2 T11645
43 47 23.9 329 2 T32115
44 47 23.9 391 2 E83151
45 47 23.9 518 2 D71288

ALIGNMENTS

RESULT 1

A53709
alpha-conotoxin Im1 - cone shell (Conus imperialis)
N:Alternate names: alpha-CTX-Im1
C:Species: Conus imperialis (Imperial cone)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A53709
F:McIntosh, J.M.; Yoshikami, D.; Mahe, E.; Nielsen, D.B.; Rivier, J.E.; Gray, W.R.; O
J. Biol. Chem. 269, 16733-16739, 1994
A:Title: A nicotinic acetylcholine receptor ligand of unique specificity, alpha-conot
A:Reference number: A53709; MUID:94266889
A:Accession: A53709
A:Molecule type: protein
A:Residues: 1-12 <MC>
A:Note: structure confirmed by chemical synthesis
C:Comment: This alpha-conotoxin, as an acetylcholine receptor inhibitor, is a postsyn
C:Superfamily: alpha-conotoxin
C:Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; postsynaptic neu
F:2-8,3-12/Disulfide bonds: #status experimental
F:12/Modified site: amidated carboxyl end (Cys) #status experimental

Query Match 32.5%; Score 64; DB 1; Length 12;
Best Local Similarity 81.8%; Pred. No. 0.043; 0; Mismatches 2; Indels 0; Gaps 0;

Qy 24 CCSDRRCRWRC 34

Db 2 CCSDRRCRWRC 12

RESULT 2

B45512

cold-regulated protein 2 - barley (fragment)

C:Species: Hordeum vulgare (barley)

C:Date: 03-Jun-1993 #sequence_revision 30-Sep-1993 #text_change 09-Sep-1997

C:Accession: B45512

R:Cattivelli, L.; Bartels, D.

Plant Physiol. 93, 1504-1510, 1990

A:Title: Molecular cloning and characterization of cold-regulated genes in barley.

A:Reference number: A45512

A:Accession: B45512

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-202 <CAT>

A:Cross-references: GB:M60733; NID:g167027; PID:g167028

Query Match

Best Local Similarity 28.9%; Score 57; DB 2; Length 202;

Matches 14; Conservative 1; Mismatches 15; Indels 0; Gaps 0;

QY 4 RNAPADDDKASDLIAQIVRRACCSDDRCRW 33
| | | | | | | : | | | | | | |
Db 91 RRAPASGAARVLRREGAERGCSDTRCRCR 120

RESULT 3
T46250
hypothetical protein DKF2p761A051.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T46250
R:Ottewaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23031
A:Accession: T46250
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-521 <AAA>
A:Cross-references: EMBL:AL137496
A:Experimental source: adult amygdala; clone DKF2p761A051
C:Genetics:
Note: DKF2p761A051.1

Query Match 28.9%; Score 57; DB 2; Length 521;
Best Local Similarity 53.3%; Pred. No. 7;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 18 QIVRRACCSDDRCRW 32
: | : | | | | | | |
Db 435 KIFEISCCSDHRCKW 449

RESULT 4
AF2828
FAD dependent oxidoreductase Atu2053 [imported] - Agrobacterium tumefaciens (strain C58,
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
C:Accession: AF2828
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AF2828
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-470 <KUR>
A:Cross-references: GB:AE008688; PIDN:AAL43044.1; PID:g17740510; GSPDB:GN00186
R:Experimental source: strain C58 (Dupont)
C:Genetics:

A:Gene: Atu2053
A:Map position: circular chromosome
C:Superfamily: glycolate oxidase chain glcd
Query Match 27.4%; Score 54; DB 2; Length 470;
Best Local Similarity 36.7%; Pred. No. 16;
Matches 11; Conservative 6; Mismatches 13; Indels 0; Gaps 0;
QY 1 FDGRNAPADDDKASDLIAQIVRRACCSDDRC 30
| | | | | | | : | | | | | | |
Db 394 FDGKDAESVAKTEAFVRLNRRRAIAMDGTC 423

RESULT 5
D97606
glycolate oxidase chain D-like, D-lactate dehydrogenase-like (AP002543) [imported] - Ag
C:Species: Agrobacterium tumefaciens
C:Date: 1-30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002

C:Accession: D97606
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorillo, B.; Goldm
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A:Reference number: A97359; PMID:11743194
A:Accession: D97606
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-477 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK87805.1; PID:g15157181; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C3718
A:Map position: circular chromosome
C:Superfamily: glycolate oxidase chain glcd

Query Match 27.4%; Score 54; DB 2; Length 477;
Best Local Similarity 36.7%; Pred. No. 16;
Matches 11; Conservative 6; Mismatches 13; Indels 0; Gaps 0;
QY 1 FDGRNAPADDDKASDLIAQIVRRACCSDDRC 30
| | | | | | | : | | | | | | |
Db 401 FDGKDAESVAKTEAFVRLNRRRAIAMDGTC 430

RESULT 6
A42325
orf 5' to pheC - Pseudomonas aeruginosa (fragment)
C:Species: Pseudomonas aeruginosa
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994
C:Accession: A42325
R:Zhao, G.S.; Xia, T.H.; Fischer, R.S.; Jensen, R.A.
J. Biol. Chem. 267, 2487-2493, 1992
A:Title: Cyclohexadienyl dehydratase from Pseudomonas aeruginosa. Molecular cloning o
A:Reference number: A42325; MUID:92129331
A:Accession: A42325
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-72 <ZHA>
A:Note: sequence extracted from NCBI backbone (NCBIN:78415, NCBIP:78416)

Query Match 26.6%; Score 52.5; DB 2; Length 72;
Best Local Similarity 45.0%; Pred. No. 5.9;
Matches 9; Conservative 1; Mismatches 3; Indels 7; Gaps 1;
QY 22 RACCSDD-----RCRWRC 34
| | | : | | | | |
Db 13 RCCCATTRNGCRGYRCRWRC 32

RESULT 7
A12372
hypothetical protein alr4537 [imported] - Anabaena sp. (strain PCC 7120)
C:Species: Anabaena sp.
A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C:Accession: A12372
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: A12372
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-252 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA076236.1; PID:g17133673; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr4537

Query Match 26.4%; Score 52; DB 2; Length 252;
Best Local Similarity 39.0%; Pred. No. 18;
Matches 16; Conservative 3; Mismatches 8; Indels 14; Gaps 3;

OY 3 GRNAPADKASDLIAQI-----VRRAC-----CSDRR 30

Db 91 GRN-PSADAVEPLIAQYRLDNGYVRKGVVWALGNCHDRRC 130

RESULT 8

AF3639 tRNA pseudouridine synthase A (EC 4.2.1.70) [imported] - Brucella melitensis (strain 16M)

C:Species: Brucella melitensis

C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002

C:Accession: AF3639

A:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujar, C.; Los, T.; Ivanova,

A: Mazur, M.; Goltsman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Letess

A: oc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A:Reference number: AD3252; PMID:11756688

A:Accession: AF3639

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-565 <KUR>

A:Cross-references: GB:AE008918; PIDN:AAL54281.1; PID:g17985257; GSPDB:GN00191

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BME11039

A:Map position: II

C:Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 26.4%; Score 52; DB 2; Length 565;
Best Local Similarity 37.9%; Pred. No. 34;
Matches 11; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

OY 1 FDGRNAPADKASDLIAQIVRRACCSDRR 29

Db 45 FCGSRGAPVEETSERIAKRLARIGASRR 73

RESULT 9

AD2812

dihydroxy-acid dehydratase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002

C:Accession: AD2812

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I

A: erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell

A: Karp, P.; Romero, P.; Zhang, S.

A: Science 294, 2317-2322, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

A: ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AD2812

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-611 <KUR>

A:Cross-references: GB:AF008688; PIDN:AAL42914.1; PID:g17740369; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: ilvD

A:Map position: circular chromosome

C:Superfamily: dihydroxy-acid dehydratase

Query Match 26.4%; Score 52; DB 2; Length 611;
Best Local Similarity 45.8%; Pred. No. 37;
Matches 11; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

OY 2 DGRNAPADKASDLIAQIVRRACC 25

Db 166 DAWAAADDKISDEDVKIISAC 189

RESULT 10

E97590

dihydroxy-acid dehydratase [imported] - Agrobacterium tumefaciens (strain C58, Cerson

C:Species: Agrobacterium tumefaciens

C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002

C:Accession: E97590

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ouello, B.; Goldm

A: A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz,

A: Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium

A:Reference number: A97359; PMID:11743194

A:Accession: E97590

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-642 <KUR>

A:Cross-references: GB:AE007869; PIDN:AAK87678.1; PID:g15157037; GSPDB:GN00169

C:Genetics:

A:Gene: AGR_C_3510

A:Map position: circular chromosome

C:Superfamily: dihydroxy-acid dehydratase

Query Match 26.4%; Score 52; DB 2; Length 642;
Best Local Similarity 45.8%; Pred. No. 38;
Matches 11; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

OY 2 DGRNAPADKASDLIAQIVRRACC 25

Db 197 DAWAAADDKISDEDVKIISAC 220

RESULT 11

JN0057

hypothetical 17K protein (pili 5' region) - Pseudomonas aeruginosa

N:Alternate names: Hypothetical protein 1

C:Species: Pseudomonas aeruginosa

C:Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 08-Oct-1999

C:Accession: JN0057

R:Whitchurch, C.B.; Hobbs, M.; Livingston, S.P.; Krishnapillai, V.; Mattick, J.S.

A:Gene 101, 33-44, 1991

A:Title: Characterisation of a Pseudomonas aeruginosa twitching motility gene and evi

A:Reference number: JN0055; PMID:91285432

A:Accession: JN0057

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-157 <WHI>

A:Cross-references: GB:M55524; NID:g151483; PIDN:AAA25960.1; PID:g151487

C:Superfamily: Pseudomonas aeruginosa hypothetical 17K protein (pili 5' region)

Query Match 26.1%; Score 51.5; DB 2; Length 157;

Best Local Similarity 27.8%; Pred. No. 15;

Matches 15; Conservative 5; Mismatches 9; Indels 25; Gaps 3;

OY 3 GRNAPADKASDLIAQIVRRAC-----CSDRR-----RWRC 34

Db 73 GRSSGAQ---PDLLASPLTLTCRTFRGGSPAGRCSDRCAIFNRSTECTHWKC 123

RESULT 12

C70842

probable pmmb protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 02-Mar-2001

C:Accession: C70842

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon

A: Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,

A: Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

A: Nature 393, 537-544, 1998

A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987
A;Accession: C70842
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-534 <COL>
A;Cross-references: GB:AL021841; GB:AL123456; NID:g3261517; PIDN:CAAL7080.1; PID:g289421
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: pmbB
C;Superfamily: Mycoplasma pirum phosphomannomutase

Query Match 26.1%; Score 51.5; DB 2; Length 534;
Best Local Similarity 45.8%; Pred. No. 38;
Matches 11; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

QY 7 PADKASDLIAQIVRRACCSDDRC 30
||||| : : |||
Db 184 PSENRASDLIDRYIQRA-AGVRC 206

SULT 13
T25228
hypothetical protein T24C2.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C;Accession: T25228
R;Baynes, C.
submitted to the EMBL Data Library, November 1995
A;Reference number: Z20000
A;Accession: T25228
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-665 <WIL>
A;Cross-references: EMBL:Z68120; PIDN:CAA92199.1; GSPDB:GN00028; CESP:T24C2.1
A;Experimental source: clone T24C2
C;Genetics:
A;Gene: CRSP:T24C2.1
A;Map position: X
A;Introns: 31/1; 57/3; 95/3; 141/2; 175/2; 196/3; 226/1; 259/1; 328/3; 388/2; 431/2; 523/2
C;Superfamily: Caenorhabditis elegans hypothetical protein T24C2.1

Query Match 26.1%; Score 51.5; DB 2; Length 665;
Best Local Similarity 36.4%; Pred. No. 46;
Matches 12; Conservative 6; Mismatches 6; Indels 9; Gaps 2;

QY 9 DDKA---SDLIAQIVRRAC-----CSDRRCRW 32
||||| : : |||
536 DDKAINDSDIVSEKQTSACYNATIISQROCKW 568

RESULT 14
D90657
hypothetical protein ECs0228 [imported] - Escherichia coli (strain O157:H7, substrain R1
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C;Accession: D90657
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: D90657
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <HAY>
A;Cross-references: PIDN:BA000007; PIDN:BA033651.1; PID:g13359684; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain R1MD 0509952
C;Genetics:
A;Gene: ECs0228

Query Match 25.9%; Score 51; DB 2; Length 360;
Best Local Similarity 45.0%; Pred. No. 33;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 5 NAPADKASDLIAQIVRRAC 24
||||| : : |||
Db 16 NPPADKEQNDPLAQVFNAC 35

RESULT 15
D85508
hypothetical protein Z0259 [imported] - Escherichia coli (strain O157:H7, substrain E
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: D85508
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
illier, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D85508
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <STO>
A;Cross-references: GB:AE005174; NID:g12512965; PIDN:AAG54528.1; GSPDB:GN00145; UWGP:
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z0259

Query Match 25.9%; Score 51; DB 2; Length 360;
Best Local Similarity 45.0%; Pred. No. 33;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 5 NAPADKASDLIAQIVRRAC 24
||||| : : |||
Db 16 NPPADKEQNDPLAQVFNAC 35

Search completed: July 1, 2002, 12:25:53
Job time: 165 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 12:27:03 ; Search time 16.34 Seconds
(without alignments)
82.937 Million cell updates/sec

Title: US-09-493-795A-236

Perfect score: 197

Sequence: 1 FDGRNAPADDKASDLIAQIVRRACCSDDRRRCWG 35

Scoring table:

BLOSUM62DX

Gapop 10.0 , Gapext 0.5

arched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 70.5 | 35.8 | 69 | 1 | CXAL_CONTE |
| 2 | 64 | 32.5 | 12 | 1 | CXAL_CONIM |
| 3 | 57.5 | 29.2 | 66 | 1 | CXAL_CONTE |
| 4 | 57 | 28.9 | 202 | 1 | CR2_HORVU |
| 5 | 51 | 25.9 | 863 | 1 | PHSG_MYCTU |
| 6 | 49.5 | 25.1 | 490 | 1 | FALD_RABIT |
| 7 | 49.5 | 25.1 | 825 | 1 | QUTA_EMENI |
| 8 | 49 | 24.9 | 88 | 1 | CR14_HORVU |
| 9 | 49 | 24.9 | 145 | 1 | PA2M_CAVPO |
| 10 | 48.5 | 24.6 | 271 | 1 | CASP_CHICK |
| 11 | 48 | 24.4 | 381 | 1 | SELP_HUMAN |
| 12 | 48 | 24.4 | 996 | 1 | ATAL_MAKNI |
| 13 | 47.5 | 24.1 | 404 | 1 | IIBC_PIG |
| 14 | 47.5 | 24.1 | 503 | 1 | IRF7_HUMAN |
| 15 | 47 | 23.9 | 110 | 1 | VL3_HPV5B |
| 16 | 47 | 23.9 | 306 | 1 | IF2A_SCHPO |
| 17 | 47 | 23.9 | 380 | 1 | PHLC_TRYCR |
| 18 | 47 | 23.9 | 380 | 1 | FUMB_ECOLI |
| 19 | 47 | 23.9 | 548 | 1 | CE1B_ECOLI |
| 20 | 47 | 23.9 | 626 | 1 | CE1B_ECOLI |
| 21 | 47 | 23.9 | 755 | 1 | COMP_RAT |
| 22 | 47 | 23.9 | 1053 | 1 | HMDH_SCHPO |
| 23 | 46.5 | 23.6 | 364 | 1 | WNT6_MOUSE |
| 24 | 46.5 | 23.6 | 365 | 1 | WNT6_HUMAN |
| 25 | 46.5 | 23.6 | 527 | 1 | GP2_HUMAN |
| 26 | 46.5 | 23.6 | 668 | 1 | TKT1_PASMU |
| 27 | 46.5 | 23.6 | 668 | 1 | TKT2_PASMU |
| 28 | 46 | 23.4 | 396 | 1 | PIRF_SORMA |
| 29 | 46 | 23.4 | 397 | 1 | PIRF_NEUCR |
| 30 | 46 | 23.4 | 541 | 1 | AAAT_RABIT |
| 31 | 46 | 23.4 | 771 | 1 | HEPA_VZVD |
| 32 | 46 | 23.4 | 1840 | 1 | SUIS_RAT |
| 33 | 45.5 | 23.1 | 407 | 1 | FA7_BOVIN |

34 45.5 23.1 501 1 BACE_HUMAN P56817 homo sapien
35 45.5 23.1 501 1 BACE_MOUSE P56818 mus musculus
36 45.5 23.1 501 1 BACE_RAT P56819 rattus norv
37 45.5 23.1 642 1 HEN1_RAT P13195 rattus norv
38 45.5 23.1 1350 1 NOS_DROME Q27571 drosophila
39 45 22.8 50 1 HSP1_PONPY P35310 pongo pygma
40 45 22.8 104 1 HSP2_RAT P11248 rattus norv
41 45 22.8 468 1 T10A_HUMAN Q00220 homo sapien
42 45 22.8 514 1 PMGI_ECOLI P37589 escherichia
43 45 22.8 555 1 FPRB_MYCLE Q33064 mycobacteri
44 45 22.8 575 1 FPRB_MYCTU Q10547 mycobacteri
45 45 22.8 577 1 FZD5_MOUSE Q9eqd0 mus musculus

ALIGNMENTS

RESULT 1
CXAL_CONTE STANDARD; PRT; 69 AA.
AC Q9XZK6:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-type conotoxin Tx1 precursor.
OS Conus textile (Cloth-of-gold cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6494;
RN [1]
RP SEQUENCE FROM N.A.
RA Lu B.S., Yu F., Huang P.T., Huang C.F.;
RT "Precursor sequences of conotoxins from Conus textile.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE ALPHA-TYPE CONOTOXIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF146352; AAD31912.1;
KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Venom;
KW Signal.
FT SIGNAL 1 21 POTENTIAL.
FT PROPEP 22 48 POTENTIAL.
FT PEPTIDE 49 69 ALPHA-TYPE CONOTOXIN TX1.
FT DISULFID 51 57 BY SIMILARITY.
FT DISULFID 52 65 BY SIMILARITY.
SQ SEQUENCE 69 AA; 7442 MW; E36ECE90BF1B56B0 CRC64;

Query Match 35.8%; Score 70.5; DB 1; Length 69;

Best Local Similarity 58.1%; Pred. No. 0.0066;

Matches 18; Conservative 2; Mismatches 8; Indels 3; Gaps 2;

QY 1 FDGRNAPADDKASDLIAQIVRR-ACCSDDRRRC 30

Db 29 FHGRNAAA--KASGLVSLTDRRPECSCDPRC 57

RESULT 2

CXAL_CONIM STANDARD; PRT; 12 AA.

AC P50983;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-conotoxin ImI.
OS Conus imperialis (Imperial cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=35631;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=Venom;
RX MEDLINE=94266889; PubMed=8206995;
RA McIntosh J.M., Yoshikami D., Mahe E., Nielsen D.B., Rivier J.E.,
RA Gray W.R., Oliveira B.M.;
RT "A nicotinic acetylcholine receptor ligand of unique specificity,
RT Alpha-conotoxin ImI."
RL J. Biol. Chem. 269:16733-16739(1994).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=95379776; PubMed=7651351;
RA Johnson D.S., Martinez J., Elgoyhen A.B., Heinemann S.F.,
RA McIntosh J.M.;
RT "Alpha-conotoxin ImI exhibits subtype-specific nicotinic
RT acetylcholine receptor blockade: preferential inhibition of homomeric
RT alpha 7 and alpha 9 receptors."
RL Mol. Pharmacol. 48:194-199(1995).
RN [3]
RP STRUCTURE BY NMR.
RX MEDLINE=99212205; PubMed=10194298;
RA Rogers J.P., Lugnbuhl P., Shen G.S., McCabe R.T., Stevens R.C.,
RA Wemmer D.E.;
RT "NMR solution structure of alpha-conotoxin ImI and comparison to
RT other conotoxins specific for neuronal nicotinic acetylcholine
RT receptors."
RL Biochemistry 38:3874-3882(1999).
RN [4]
RP STRUCTURE BY NMR.
RX MEDLINE=99158061; PubMed=10050774;
RA Maslennikov I.V., Shenkarev Z.O., Zhmak M.N., Ivanov V.T.,
RA Methfessel C., Tsetlin V.I., Arseniev A.S.;
RT "NMR spatial structure of alpha-conotoxin ImI reveals a common
RT scaffold in snail and snake toxins recognizing neuronal nicotinic
RT acetylcholine receptors."
RL FEBS Lett. 444:275-280(1999).
RN [5]
RP STRUCTURE BY NMR.
RX MEDLINE=99324017; PubMed=10395477;
RA Gehrmann J., Daly N.L., Alewood P.F., Craik D.J.;
RT "Solution structure of alpha-conotoxin ImI by 1H nuclear magnetic
RT resonance."
RL J. Med. Chem. 42:2364-2372(1999).
RN [6]
RP -1- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM. IT IS HIGHLY ACTIVE AGAINST THE NEUROMUSCULAR
CC RECEPTOR IN FROG BUT NOT IN MICE. IN CONTRAST, IT INDUCES SEIZURES
CC WHEN INJECTED CENTRALLY IN MICE AND RATS. IT TARGET NEURONAL
CC NACHRS IN MAMMALS. BLOCKS HOMOMERIC ALPHA-7 NICOTINIC RECEPTORS
CC WITH THE HIGHEST APPARENT AFFINITY AND HOMOMERIC ALPHA-9 RECEPTORS
CC WITH 8-FOLD LOWER AFFINITY. HAS NO EFFECT ON RECEPTORS COMPOSED OF
CC ALPHA-2/BETA-2, ALPHA-3/BETA-2, ALPHA-4/BETA-2, ALPHA-2/BETA-4,
CC ALPHA-3/BETA-4, OR ALPHA-4/BETA-4 SUBUNIT COMBINATIONS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE ALPHA-TYPE CONOTOXIN FAMILY.
DR PDB; 1MI; 15-JUN-99.
DR PDB; 1MI; 23-APR-99.
DR PDB; 1CN; 27-MAY-99.
KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
KW Venom; 3D-structure.
FT DISULFID 2 8
FT DISULFID 3 12
FT MOD_RES 12 12
FT SEQUENCE 12 AA; 1357 MW; 9C29CEA545A4176A CRC64;
AMIDATION

Query Match 32.5%; Score 64; DB 1; Length 12;
Best Local Similarity 81.8%; Pred. No. 0.011;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 24 CCSDRRRCRWRC 34
DB 2 CCSDPRCAWRC 12
|||||
RESULT 3
CXAZ2_CONTE STANDARD; PRT; 66 AA.
AC Q9XZK7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-type conotoxin Tx2 precursor.
OS Conus textile (Cloth-of-gold cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6494;
RN [1]
RP SEQUENCE FROM N.A.
RA Lu B.S., Yu F., Huang P.T., Huang C.F.;
RT "Precursor sequences of conotoxins from Conus textile.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE ALPHA-TYPE CONOTOXIN FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF146353; AAD31913.1; -;
KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Venom;
KW Signal.
FT SIGNAL 1 21 POTENTIAL.
FT PROPEP 22 48 POTENTIAL.
FT PEPTIDE 49 66 ALPHA-TYPE CONOTOXIN TX2.
FT DISULFID 51 57 BY SIMILARITY.
FT DISULFID 52 65 BY SIMILARITY.
SQ SEQUENCE 66 AA; 7254 MW; EDD59BBAB94F26F CRC64;
Query Match 29.2%; Score 57.5; DB 1; Length 66;
Best Local Similarity 51.6%; Pred. No. 0.34;
Matches 16; Conservative 2; Mismatches 10; Indels 3; Gaps 2;
QY 1 FDGNAPADKASDLIAQIVRR-ACCSDRRC 30
DB 29 FHGRNAAA--KASGLVSLTDRRPECCSHPC 57
|||||
RESULT 4
CR2_HORVU STANDARD; PRT; 202 AA.
AC P23252;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-NOV-1991 (Rel. 20, Last annotation update)
DE Cold-regulated protein 2 (Fragment).
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Poales; Poaceae; Poideae;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poaceae; Hordeum.
OX NCBI_TaxID=4513;

| MOD_RES | 72 | 72 | GAMMA-CARBOXYGLUTAMIC ACID (BY SIMILARITY). |
|--|-----------------|----------------|---|
| MOD_RES | 75 | 75 | GAMMA-CARBOXYGLUTAMIC ACID (BY SIMILARITY). |
| MOD_RES | 79 | 79 | GAMMA-CARBOXYGLUTAMIC ACID (BY SIMILARITY). |
| MOD_RES | 103 | 103 | HYDROXYLATION (BY SIMILARITY). |
| ACT_SITE | 274 | 274 | CHARGE RELAY SYSTEM. |
| ACT_SITE | 320 | 320 | CHARGE RELAY SYSTEM. |
| ACT_SITE | 417 | 417 | CHARGE RELAY SYSTEM. |
| DISULFID | 90 | 101 | BY SIMILARITY. |
| DISULFID | 95 | 110 | BY SIMILARITY. |
| DISULFID | 112 | 121 | BY SIMILARITY. |
| DISULFID | 129 | 140 | BY SIMILARITY. |
| DISULFID | 136 | 149 | BY SIMILARITY. |
| DISULFID | 151 | 164 | BY SIMILARITY. |
| DISULFID | 172 | 340 | INTERCHAIN (BY SIMILARITY). |
| DISULFID | 239 | 244 | BY SIMILARITY. |
| DISULFID | 259 | 275 | BY SIMILARITY. |
| DISULFID | 388 | 402 | BY SIMILARITY. |
| DISULFID | 413 | 441 | BY SIMILARITY. |
| CARBOHYD | 61 | 61 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| CARBOHYD | 187 | 187 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| CARBOHYD | 205 | 205 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| SEQUENCE | 490 AA; | 53965 MW; | 3A39FA85AF2A6D11 CRC64; |
| Query Match | 25.1%; | Score 49.5; | DB 1; Length 490; |
| Best Local Similarity | 34.4%; | Pred. No.23; | 15; Gaps 1; |
| Matches 11; | Conservative 5; | Mismatches 15; | Indels 1; |
| 2 DGRNAPADKASDLIAQIVRRACCSDDRRGRWR 33 | | | |
| 217 DSPEPPEDSSSLV-RIVGGQDRGECFPWQ 247 | | | |
| QUOTA_EMENI | STANDARD; | PRT; | 825 AA. |
| P10563; | | | |
| 01-JUL-1989 (Rel. 11, Created) | | | |
| 01-JUL-1989 (Rel. 11, Last sequence update) | | | |
| 01-APR-1993 (Rel. 25, Last annotation update) | | | |
| Quinic acid utilization activator. | | | |
| QuaT | | | |
| Emicella nidulans (Aspergillus nidulans). | | | |
| Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; | | | |
| Eurotiales; Trichocomaceae; Emicella. | | | |
| NCBI_Taxid=5072; | | | |
| SEQUENCE FROM N.A. | | | |
| MEDLINE=88040423; PubMed=3313276; | | | |
| Beri R.K., Whittington H., Roberts C.F., Hawkins A.R.; | | | |
| "Isolation and characterization of the positively acting regulatory gene QUTa from Aspergillus nidulans." | | | |
| Nucleic Acids Res. 15:7991-8001(1987). | | | |
| FUNCTION: TRANSCRIPTION ACTIVATION OF GENES FOR ENZYMES AND PROTEINS OF QUINATE METABOLISM BY BINDING TO A 16 BASE-PAIR SEQUENCE (CONSENSUS GGATAANNNTATCC) IN FRONT OF EACH QUT GENE. | | | |
| SUBCELLULAR LOCATION: Nuclear. | | | |
| SIMILARITY: CONTAINS 1 ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR CLUSTER DOMAIN. | | | |
| This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch). | | | |
| EMBL; X06252; CAA25994.1; | | | |
| PIR; A26983; A26983. | | | |


```
Matches 13; Conservative 3; Mismatches 13; Indels 4; Gaps 2;
QY 3 GRNPADKASDLIAQIVRRACCSRRRCRCG 35
   | | | | | | | | | | | | | | | |
Db 52 GKCTPKD--ATDRCC--VRHDCCYDLRMKRGCG 80
RESULT 10
CASP_CHICK STANDARD; PRT; 271 AA.
AC Q90830;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Cartilage associated protein precursor (Dualin).
GN CRTAP OR CASP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
SEQUENCE FROM N.A.
RX MEDLINE=97360293; PubMed=9217321;
RA Castagnola P., Gennari M., Morello R., Tonachini L., Marin O.,
RA Gaggero A., Cancedda R.;
RT "Cartilage associated protein (CASP) is a novel developmentally
RT regulated chick embryo protein.";
RL J. Cell Sci. 110:1351-1359(1997).
CC -!- SUBCELLULAR LOCATION: Secreted. Extracellular matrix.
CC -!- TISSUE SPECIFICITY: Found in articular chondrocytes. Expressed in
CC a variety of tissues.
CC -!- SIMILARITY: BELONGS TO THE CRTAP / NO55 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X97607; CAA66206.1; -.
KW Extracellular matrix; Signal.
FT SIGNAL 1 15 POTENTIAL.
FT CHAIN 16 271 CARTILAGE ASSOCIATED PROTEIN.
FT CARBOHYD 76 76 N-LINKED (GLCNAC...)(POTENTIAL).
FT SEQUENCE 271 AA; 31493 MW; A56343DBF604914C CRC64;
Query Match 24.6%; Score 48.5; DB 1; Length 271;
Best Local Similarity 44.4%; Pred. No. 19;
Matches 12; Conservative 7; Mismatches 5; Indels 3; Gaps 2;
QY 7 PADKASD--LIAQIVRRACCSRRRCRC 31
   | | | | | | | | | | | | | |
Db 88 PADGELAEELRLAGVLRRAQCL-RRCK 113
RESULT 11
SELP_HUMAN STANDARD; PRT; 381 AA.
AC P49908;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Selenoprotein P precursor (Sep).
GN SEPP1 OR SELP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
```

```
SEQUENCE FROM N.A.
RP TISSUE=Liver, and Heart;
RX MEDLINE=93133823; PubMed=8421687;
RA Hill K.E., Lloyd R.S., Burk R.F.;
RT "Conserved nucleotide sequences in the open reading frame and 3'
RT untranslated region of selenoprotein P mRNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:537-541(1993).
RN [2]
RP PARTIAL SEQUENCE.
RX MEDLINE=94191007; PubMed=8142465;
RA Aakesson B., Belliew T., Burk R.F.;
RT "Purification of selenoprotein P from human plasma.";
RL Biochim. Biophys. Acta 1204:243-249(1994).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=98413836; PubMed=9735174;
RA Mostert V., Lombeck I., Abel J.;
RT "A novel method for the purification of selenoprotein P from human
RT plasma.";
RL Arch. Biochem. Biophys. 357:326-330(1998).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=20239644; PubMed=10775431;
RA Mostert V.;
RT "Selenoprotein P: properties, functions, and regulation.";
RL Arch. Biochem. Biophys. 376:433-438(2000).
RN [5]
RP REVIEW.
RX MEDLINE=95017128; PubMed=7931697;
RA Burk R.F., Hill K.E.;
RT "Selenoprotein P. A selenium-rich extracellular glycoprotein.";
RL J. Nutr. 124:1891-1897(1994).
CC -!- FUNCTION: MIGHT BE RESPONSIBLE FOR SOME OF THE EXTRACELLULAR
CC ANTIOXIDANT DEFENSE PROPERTIES OF SELENIUM OR MIGHT BE INVOLVED IN
CC THE TRANSPORT OF SELENIUM.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MADE IN THE LIVER AND HEART AND SECRETED INTO
CC THE PLASMA. IT IS ALSO FOUND IN THE KIDNEY.
CC -!- MISCELLANEOUS: THE SELENOCYSTEINES ARE ALL ENCODED BY THE OPAL
CC CODON, UGA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z11793; CAA77836.1; -.
KW MIM; 601484; -.
KW Glycoprotein; Signal; Selenium; Selenocysteine; Plasma.
FT SIGNAL 1 19 SELENOPROTEIN P.
FT CHAIN 20 381 POLY-HIS.
FT DOMAIN 244 249
FT SE_CYS 59 59
FT SE_CYS 300 300
FT SE_CYS 318 318
FT SE_CYS 330 330
FT SE_CYS 345 345
FT SE_CYS 352 352
FT SE_CYS 367 367
FT SE_CYS 369 369
FT SE_CYS 376 376
FT SE_CYS 378 378
FT CARBOHYD 46 46 N-LINKED (GLCNAC...)(POTENTIAL).
FT CARBOHYD 83 83 N-LINKED (GLCNAC...)(POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC...)(POTENTIAL).
FT CARBOHYD 128 128 N-LINKED (GLCNAC...)(POTENTIAL).
FT CARBOHYD 174 174 N-LINKED (GLCNAC...)(POTENTIAL).
FT CARBOHYD 338 338 N-LINKED (GLCNAC...)(POTENTIAL).
FT SEQUENCE 381 AA; 42705 MW; 1A82FFA0BCD13CD7 CRC64;
SQ
```


| | | | |
|------------------|--|--|--------------------------------|
| Qy | 3 | GRNPADDKASDLIAQIVRR---AC--CSDRRR | 30 |
| I | | | ::: |
| D | b | 45 | GENATVMDKARALIDSVIRKPQACICINHC |
| | | | 77 |
| | | | |
| RESULT | 14 | | |
| IRF7_HUMAN | | | |
| ID | IRF7_HUMAN | STANDARD; | PRT; 503 AA. |
| AC | Q929B5; | O00332; | O00333; O75924; |
| DT | 01-NOV-1997 | (Rel. 35, | Created) |
| DD | 01-NOV-1997 | (Rel. 35, | Last sequence update) |
| DT | 16-OCT-2001 | (Rel. 40, | Last annotation update) |
| DE | Interferon regulatory factor 7 (IRF-7). | | |
| GN | IRF7. | | |
| OS | Homo sapiens (Human). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| CC | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | |
| NCBI_TaxID=9606; | [1] | | |
| RN | SEQUENCE FROM N.A. (ISOFORM A). | | |
| RP | TISSUE=Splice; | | |
| RC | Grossman A., Nicholl J., Antonio L., Luethy R., Suggs S., | | |
| RA | Sutherland G.R., Mak T.W.; | | |
| RL | Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases. | | |
| RN | [2] | | |
| RP | SEQUENCE FROM N.A. (ISOFORMS A; B AND C). | | |
| RX | MEDLINE=97459673; PubMed=9315633; | | |
| RT | Zhang L., Pagano J.S.; | | |
| RF | "IRF-7, a new interferon regulatory factor associated with Epstein- | | |
| RT | Barr virus latency."; | | |
| RL | Mol. Cell. Biol. 17:5748-5757(1997). | | |
| RN | [3] | | |
| RP | SEQUENCE FROM N.A. (ISOFORM D). | | |
| RX | MEDLINE=99003279; PubMed=9786932; | | |
| RT | Au W.-C., Moore P.A., Lafleur D.W., Tombal B., Pittha P.M.; | | |
| NT | "Characterization of the interferon regulatory factor-7 and its | | |
| RT | potential role in the transcription activation of Interferon A | | |
| RT | genes."; | | |
| RL | J. Biol. Chem. 273:29210-29217(1998). | | |
| CC | -1- FUNCTION: TRANSCRIPTIONAL ACTIVATOR. BINDS TO THE INTERFERON- | | |
| CC | STIMULATED RESPONSE ELEMENT (ISRE) IN IFN PROMOTERS AND IN THE Q | | |
| CC | PROMOTOR (QP) OF EBV NUCLEAR ANTIGEN-1 (EBNA1). | | |
| CC | -1- SUBCELLULAR LOCATION: Nuclear (Potential). | | |
| CC | -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS: A (SHOWN HERE), B/BETA, C/GAMMA | | |
| CC | AND D/H. ARE PRODUCED BY ALTERNATIVE SPLICING. | | |
| CC | TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN SPLEEN, THYMUS, AND | | |
| CC | PERIPHERAL BLOOD LEUKOCYTES. | | |
| CC | -1- SIMILARITY: BELONGS TO THE IRF FAMILY. | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration - | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | |
| CC | use by non-profit institutions as long as its content is in no way | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | | |
| CC | or send an email to license@isb-sib.ch). | | |
| EMBL; U73036; | AAB17190.1; | - | |
| DR | EMBL; U53830; | AAB80686.1; | - |
| DR | EMBL; U53831; | AAB80688.1; | - |
| DR | EMBL; U53832; | AAB80690.1; | - |
| DR | EMBL; AF076494; | AAC70999.1; | - |
| DR | HSP; P23906; | 2IRF. | |
| MIM; 605047; | -- | | |
| DR | InterPro; IPR001346; | IRF. | |
| DR | pfam; PF00605; | IRF. 1. | |
| DR | PRINTS; PR00267; | INTFRNREGFC. | |
| DR | ProDom; PD002355; | IRF. 1. | |
| DR | SMART; SM00348; | IRF. 1. | |
| DR | PROSITE; PS00601; | IRF. 1. | |
| KW | Transcription regulation; | DNA-binding; Nuclear protein; Activator; | |
| KW | Alternative splicing. | | |

Query Match 24.1%; Score 47.5; DB 1; Length 404;
Best Local Similarity 39.4%; Pred. No. 36;
Matches 13; Conservative 4; Mismatches 11; Indels 5; Gaps 2;

FT DNA_BIND 13 122 TRYPTOPHAN PENTAD REPEAT.
FT VARSPLIC 1 6 MALAPE -> MPVPERPAAGPDSRPGR (IN ISOFORM
D).
FT VARSPLIC 152 164 GGPGPFLATHA -> AQSLLGSGCTGG (IN
ISOFORM C).
FT VARSPLIC 165 503 MISSING (IN ISOFORM C).
FT VARSPLIC 228 256 MISSING (IN ISOFORM B).
FT CONFLICT 179 179 E -> K (IN REF. 2).
FT CONFLICT 412 412 Q -> R (IN REF. 3).
SQ SEQUENCE 503 AA; 54278 MW; AA6A39E0E272727C CRC64;

Query Match 24.1%; Score 47.5; DB 1; Length 503;
Best Local Similarity 45.5%; Pred. No. 44;
Matches 10; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

Oy 6 APADKASDLIAQIVRRACCS 27
174 APAGDE-GDLLLQAVQSC 194

RESULT 15
VL3_HPV5B STANDARD; PRT; 110 AA.
AC P26541;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Probable L3 protein.
OS Human papillomavirus type 5b.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10599;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91306467; PubMed=1649510;
RA Yabe Y., Sakai A., Hitsumoto T., Kato H., Ogura H.;
RT "A subtype of human papillomavirus 5 (HPV-5b) and its subgenomic
segment amplified in a carcinoma: nucleotide sequences and genomic
organizations.";
RT Virology 183:793-798(1991).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; D90252; BRA14299.1; -.
DR PIR; I40480; P3MLB5.
KW Late protein.
SQ SEQUENCE 110 AA; 13103 MW; 1B5B33A991B2510A CRC64;

Query Match 23.9%; Score 47; DB 1; Length 110;
Best Local Similarity 42.1%; Pred. No. 13;
Matches 8; Conservative 2; Mismatches 7; Indels 2; Gaps 1;

Oy 17 AQIVRRACCSRRRCWRCG 35
Db 55 ANVKENIC--DLHCRWQCG 71

Search completed: July 1, 2002, 12:27:04
Job time: 186 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 12:26:40 ; Search time 41.46 Seconds
(without alignments)
146.040 Million cell updates/sec

Title: US-09-493-795A-236
Perfect score: 197
Sequence: 1 FDGRNAPADDKASDLTAQIVERRACCSRRRCRWRCG 35

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

```
Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222
```

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000.
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database :
SPTRMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_tvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*
```

SUMMARIES

| Result No. | Score | Query % | | Length | DB | ID | Description |
|------------|-------|---------|------|--------|--------|--------|--------------|
| | | Match | | | | | |
| 1 | 74 | 37.6 | 62 | 5 | Q9U619 | Q9U619 | conus imper |
| 2 | 58 | 29.4 | 335 | 2 | Q9X5V8 | Q9X5V8 | streptomyce |
| 3 | 58 | 20.4 | 2809 | 4 | Q96JD8 | Q96JD8 | homo sapien |
| 4 | 57 | 28.9 | 428 | 4 | Q9HAU7 | Q9HAU7 | homo sapien |
| 5 | 57 | 28.9 | 521 | 4 | Q9NT71 | Q9NT71 | homo sapien |
| 6 | 57 | 28.9 | 523 | 4 | Q9HAT2 | Q9HAT2 | homo sapien |
| 7 | 57 | 28.9 | 612 | 16 | Q92M28 | Q92M28 | rhizobium m |
| 8 | 56.5 | 28.7 | 541 | 11 | P70665 | P70665 | mus musculus |
| 9 | 56.5 | 28.7 | 541 | 11 | Q922L0 | Q922L0 | mus musculus |
| 10 | 54.5 | 27.7 | 44 | 15 | O12166 | O12166 | human immun |
| 11 | 54.5 | 27.7 | 1011 | 17 | Q981K8 | Q981K8 | rhizobium l |
| 12 | 54 | 27.4 | 294 | 5 | Q96440 | Q96440 | leishmania |
| 13 | 52 | 26.4 | 314 | 2 | Q9LA88 | Q9LA88 | aeromonas h |
| 14 | 52 | 26.4 | 502 | 6 | Q9BGNG | Q9BGNG | tachyglossu |
| 15 | 51.5 | 26.1 | 305 | 2 | Q9RDP5 | Q9RDP5 | streptomyce |
| 16 | 51.5 | 26.1 | 534 | 16 | O53360 | O53360 | mycobacteri |

Q9hvv4 caenorhabdi
Q22724 caenorhabdi
Q9v1u9 drosophila
Q98b28 rhizobium l
Q946a3 theobroma c
Q9xc9 arabidopsis
Q9mb44 oryza sativ
Q94370 arabidopsis
Q92vc2 rhizobium m
Q97630 marek's dis
Q9h7j2 homo sapient
Q50330 escherichia
Q9b56 coccus penna
Q9fu06 oryza sativ
Q9rt08 deinococcus
Q68892 aquifex aeo
Q9f7f9 clostridium
Q945c8 zea mays (m
Q24136 nicotiana t
Q9vdm9 drosophila
Q9llk1 streptomyce
Q9x113 arabidopsis
Q70011 streptomyce
Q39275 equine herp
Q9u3x7 caenorhabdi
Q19085 sus scrofa
Q9cyl8 mus musculu
Q9k399 streptomyce
Q94h06 oryza sativ

ALIGNMENTS

| | |
|--------|--|
| RESULT | 1 |
| Q9U619 | |
| ID | PRELIMINARY; PRT; 62 AA. |
| AC | Q9U619; |
| DC | Q9U619; |
| DT | 01-MAY-2000 (TrEMBLrel. 13, Created) |
| DT | 01-MAY-2000 (TrEMBLrel. 13, Last sequence update) |
| DT | 01-DEC-2001 (TrEMBLrel. 19, Last annotation update) |
| DE | CONTOXIN IMIIA. |
| OE | IMIIA. |
| OS | Conus imperialis (Imperial cone). |
| GN | Conus imperialis |
| OC | Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda; |
| OC | Neogastropoda; Conoidea; Conidae; Conus. |
| OX | NCBI_Taxid=35631; |
| RN | [1] |
| RP | SEQUENCE FROM N.A. |
| RP | Zhao D., Huang P; |
| RA | "Conus imperialis conotoxin ImIIA precursor mRNA."; |
| RL | Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases. |
| RQ | EMBL; AF200595; AAF12824.1; - |
| SQ | SEQUENCE 62 AA; 6830 MW; BF0D811758C3047D CRC64; |

| | |
|----|---|
| GN | GRSB. |
| OS | Streptomyces griseus. |
| OC | Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; |
| OC | Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces |
| OX | NCBI_TaxID=1911; |
| | [1] |
| RN | SEQUENCE FROM N.A. |
| RC | STRAIN=ATCC10137; |
| RP | |
| RA | Hyun C.G., Bang J.H., Kim J.W., Han J.J., Choi Y.N., Suh J.W.; |
| RT | "Molecular cloning of the genes for deoxyhexose biosynthesis from |
| RT | Streptomyces griseus."; |
| RL | Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases. |
| DR | EMBL; AF128273; AAD31798.1; - |
| SQ | SEQUENCE 335 AA; 34490 MW; AF461DA4C90585F3 CRC64; |
| SD | |

| | | | | |
|-----------------------|--------------|----------------|----------------|-------------|
| Query Match | 29.4% | Score 58; | DB 2; | Length 335; |
| Best Local Similarity | 39.4%; | Pred. No. 2.6; | | |
| Matches 13; | Conservative | 4; | Mismatches 16; | Indels 0; |
| Matches | | | | Gaps 0; |

QY 3 GRNAPADDKASDLIAQIVRRACSDRRRCWRG 35
|| ||| | :| :||| : | |
Db 187 GRGAPAGRKDRGVLAAMVRRALAGEPLTMWHDG 219

RESULT 3
Q96JP8
ID Q96JP8
PRELIMINARY;
PRT; 2809 AA.

| | | |
|----|--|---|
| AC | 01-DEC-2001 | (TREMBLRel. 19, Created) |
| RC | 01-DEC-2001 | (TREMBLRel. 19, Last annotation update) |
| DT | 01-DEC-2001 | (TREMBLRel. 19, Last annotation update) |
| DT | 01-DEC-2001 | (TREMBLRel. 19, Last annotation update) |
| DE | FIBRILLIN3. | |
| GN | KIAA1776. | |
| GN | KIAA1776. | |
| OS | Homo sapiens (Human). | |
| OC | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | |
| OX | NCBI_TaxID=9606; | |
| OX | NCBI_TaxID=9606; | |
| RP | [1] | |
| RP | SEQUENCE FROM N.A. | |
| RP | SEQUENCE FROM N.A. | |
| RC | TISSUE=BRAIN; | |
| RC | TISSUE=BRAIN; | |
| RC | MEDLINE=21245130; PubMed=11347906; | |
| RA | Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.; | |
| RA | Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.; | |
| RT | 'Prediction of the coding sequences of unidentified human genes. XX. | |
| RT | 'Prediction of the coding sequences of unidentified human genes. XX. | |
| RT | The complete sequences of 100 new cDNA clones from brain which code | |
| RT | for large proteins in vitro. #; | |
| RL | DNA Res. 8:85-95(2001). | |
| RL | DNA Res. 8:85-95(2001). | |
| RL | EMBL; AB053450; BAB47408.1. | |
| DR | EMBL; AB053450; BAB47408.1. | |
| DR | SEQUENCE 2809 AA; 300323 MW; 20C04CC0C06C016F CRC64; | |
| SQ | SEQUENCE 2809 AA; 300323 MW; 20C04CC0C06C016F CRC64; | |

Query Match 29.4%; Score 58; DB 4; Length 2809;
Best Local Similarity 52.0%; Pred. No. 24;
Matches 13; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY 12 ASDLIAQIVRRACCSDR-RCRWRCG 35
 + + + + + + + + + + +
Db 305 AGDLAGHYTRRQCCTDGRGC-WAAG 328

| RESULT | 4 | |
|--------|---|---|
| Q9HAU7 | | |
| ID | Q9HAU7 | |
| AC | Q9HAU7: | |
| DT | 01-WAR-2001 | PRELIMINARY; PRT; 488 AA. |
| DT | 01-WAR-2001 | (TReMBLrel. 16, Created) |
| DT | 01-WAR-2001 | (TReMBLrel. 16, Last sequence update) |
| DT | 01-WAR-2001 | (TReMBLrel. 16, Last annotation update) |
| DE | SIALIC ACID-SPECIFIC 9-O-ACETYLESTERASE I. | |
| OS | Homo sapiens (Human). | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. | |
| OX | NCBI_TaxID=9606; | |
| RN | [1] | |

RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Zhu H., Li J.M., Sha J.H.;
RT "A novel gene of sialic acid-specific 9-O-acetyltransferase I from human
RL testis.";
RE Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF300796; AAG15386.1; -;
SQ SEQUENCE 488 AA; 54572 MW; DB030C82DA44916 CRC64;

Query Match 28.9%; Score 57; DB 4; Length 488;
Best Local Similarity 53.3%; Pred. No. 5.3;
Matches 8; Conservative 3; Mismatches 4; Indels

QY 18 QIVRRACSDRRCRW 32
: | : | | | | | : |
Db 402 KIFEISCCSDHRCKW 416

```

RESULT      5
Q9NT71     PRELIMINARY;          PRT;    521 AA.
AC         Q9NT71;
DT         01-OCT-2000 (TrEMBLrel. 15, Created)
DT         01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT         01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE         DE HYPOTHEtical 58.1 KDA PROTEIN (FRAGMENT).
GN         DKFZF61A051.
GS         Homo sapiens (Human).
OC         OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC         Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX         NCBI_TaxID=9606;
RN         [1]
RP         SEQUENCE FROM N.A.
RC         TISSUE=AMYGDALA;
RA         Ottenwaelder B., Obermaier B., Mewes H.W., Gassenhuber J., Wieman
RL         Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases..
DR         EMBL; ALI37496; CAB70771.1; -.
KW         Hypothetical protein.
FT         UNK_TER           1
SQ         SEQUENCE       521 AA; 58084 MW; B732FCBADF2FED8 CRC64;

```

Query Match 28.9%; Score 57; DB 4; Length 521;
Best Local Similarity 53.3%; Pred. No. 5.7;
Matches 8; Conservative 3; Mismatches 4; Indels

```
QY      18 QIVRACCSDRRCRW 32
      :| :||| | |:|
Db     435 KIFEISCCSDHRCKW 449
```

| | |
|--------|--|
| RESULT | 6 |
| Q9HAT2 | |
| ID | Q9HAT2 PRELIMINARY; PRT; 523 AA. |
| AC | O9HAT2; |
| DT | 01-MAR-2001 (TEMBLrel. 16, Created) |
| DT | 01-MAR-2001 (TEMBLrel. 16, Last sequence update) |
| DT | 01-DEC-2001 (TEMBLrel. 19, Last annotation update) |
| DE | SIALIC ACID-SPECIFIC ACETYLESTERASE II. |
| OS | Homo sapiens (human). |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. |
| CX | NCBI_TaxID=9606; |
| RN | [1] |
| RP | SEQUENCE FROM N.A. |
| RC | TISSUE=TESTIS; |
| RA | Zhu H., Zhou Z.M., Sha J.H.; |
| RT | "A novel gene of human sialic acid-specific 9-O-acetylerase from testis."; |
| RL | Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases. |
| DR | EMBL; AF303378; AAG14897.1; " |
| SO | SEQUENCE 523 AA; 58315 MW; B72CF69636D8FED8 CRC64; |

Query Match 28.98; Score 57; DB 4; Length 523;
 Best Local Similarity 53.38; Pred. No. 5.7;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 18 QIVRRACSDRRRCRW 32
 DB 437 KIFEISCCSDHRCWK 451

RESULT 7

O92M28 PRELIMINARY; PRT; 612 AA.
 AC Q92M28
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE PROBABLE DIHYDROXY-ACID DEHYDRATASE PROTEIN (EC 4.2.1.9).
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21368234; PubMed=11474104;
 RA Galibert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe F.,
 Barloy-Hubler F., Barnett M.J., Becker A., Boistard P., Bothe G.,
 Boutry L., Buhmester J., Cadieu E., Capela D., Chain P.,
 Cowie A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F.,
 Gloux S., Godrie T., Goffeau A., Goldring B., Gouzy J., Gurjal M.,
 Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,
 Kahn M.L., Kaiman S., Keating D.H., Kiss E., Komp C., Lelaure V.,
 Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelie D., Purnelle B.,
 Ramsperger U., Surzycki R., Thebault P., Vandenbol M.,
 Vorhoefer F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;
 RT "The composite genome of the legume symbiont Sinorhizobium meliloti";
 RL Science 293:668-672(2001).
 DR EMBL; ALS91792; CAC47407.1; -;
 KW Lyase; Complete proteome.
 SQ SEQUENCE 612 AA; 65468 MW; C368C51F5FC82F7 CRC64;

Query Match 28.98; Score 57; DB 16; Length 612;
 Best Local Similarity 50.08; Pred. No. 6.8;
 Matches 12; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

OY 2 DGRNAPADKASDLIAQIVRRACC 25
 DB 166 DAMVAAADKVSDEVDQIERSAC 189

RESULT 8

ID P70665 PRELIMINARY; PRT; 541 AA.
 AC P70665; Q61044;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE SIALATE O-ACETYLESTERASE PRECURSOR (EC 3.1.1.53) (SIALIC ACID-SPECIFIC
 9-O-ACETYLESTERASE) (YOLK SAC PROTEIN 2).
 GN YSG2.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (LYSOSOMAL ISOFORM).
 RC STRAIN=DBA/2 X C57BL/6;
 RX MEDLINE=97078679; PubMed=8918804;
 RA Stoddart A., Zhang Y., Paige C.J.;
 RT "Molecular cloning of the cDNA encoding a murine sialic acid-specific

9-O-acetylesterase and RNA expression in cells of hematopoietic and
 non-hematopoietic origin";
 RL Nucleic Acids Res. 24:4003-4008(1996).
 RN [2]
 RP SEQUENCE FROM N.A. (LYSOSOMAL ISOFORM).
 RX MEDLINE=96278803; PubMed=8662838;
 RA Guimaraes M.J., Bazan J.F., Castagnola J., Diaz S., Copeland N.G.,
 Gilbert D.J., Jenkins N.A., Varki A., Zlotnik A.;
 RT "Molecular Cloning and Characterization of lysosomal sialic acid O-
 acetylesterase";
 RL J. Biol. Chem. 271:13697-13705(1996).
 RN [3]
 RP SEQUENCE FROM N.A. (CYTOSOLIC ISOFORM).
 RC TISSUE=B-CELL;
 RX MEDLINE=99395134; PubMed=10464298;
 RA Takematsu H., Diaz S., Stoddart A., Zhang Y., Varki A.;
 RT "Lysosomal and cytosolic sialic acid 9-O-acetylesterase activities can
 be encoded by one gene via differential usage of a signal peptide-
 encoding exon at the N terminus";
 RL J. Biol. Chem. 274:25623-25631(1999).
 CC -!- FUNCTION: CATALYZES THE REMOVAL OF O-ACETYL ESTER GROUPS FROM
 POSITION 9 OF THE PARENT SIALIC ACID, N-ACETYLNEURAMINIC ACID.
 CC -!- CATALYTIC ACTIVITY: N-ACETYL-O-ACETYLNEURAMINATE + H(2)O = N-
 ACETYLNEURAMINATE + ACETATE.
 CC -!- ENZYME REGULATION: INHIBITED BY DIISOPROPYL FLUOROPHOSPHATE AND
 DIETHYL-P-NITROPHENYL PHOSPHATE (BY SIMILARITY).
 CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A SMALL SUBUNIT AND A
 LARGE SUBUNIT (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: LYSOSOMAL AND CYTOPLASMIC.
 CC -!- ALTERNATIVE PRODUCTS: TWO ISOFORMS; LYSOSOMAL ISOFORM (SHOWN HERE)
 AND CYTOSOLIC ISOFORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- TISSUE SPECIFICITY: THE LYSOSOMAL ISOFORM IS WIDELY EXPRESSED. THE
 CYTOSOLIC ISOFORM SHOWS A MORE RESTRICTED DISTRIBUTION WITH
 HIGHEST EXPRESSION IN BRAIN AND OVARY AND LOWER LEVELS IN LIVER
 AND THYMUS.
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN MATURE B-CELL LINES BUT NOT IN
 LESS MATURE LINES.
 CC -!- PTM: THE TWO SUBUNITS ARE DERIVED FROM A SINGLE PRECURSOR BY
 PROTEOLYTIC CLEAVAGE (BY SIMILARITY).
 CC -!- PTM: THE LYSOSOMAL ISOFORM IS GLYCOSYLATED.
 DR EMBL; U61183; AAC52880.1; -;
 DR EMBL; X98625; CAA67214.1; -;
 DR EMBL; U40408; AAB07813.1; -;
 DR EMBL; AF156856; AAD55976.1; -;
 DR MGD; MGI:104803; Ysg2.
 KW Hydrolase; Serine esterase; Glycoprotein; Lysosome; Signal;
 Alternative splicing.
 FT SIGNAL 1 23
 FT CHAIN 24 275
 FT CHAIN 276 541
 FT CARBOHYD 107 107
 FT CARBOHYD 138 138
 FT CARBOHYD 188 188
 FT CARBOHYD 293 293
 FT CARBOHYD 356 356
 FT CARBOHYD 427 427
 FT CARBOHYD 448 448
 FT CARBOHYD 462 462
 FT VARSPPLIC 1 98
 FT CONFLICT 160 160
 FT CONFLICT I -> T (IN REF. 2).
 SQ SEQUENCE 541 AA; 60775 MW; 944936C45C4A2E6B CRC64;

Query Match 28.78; Score 56.5; DB 11; Length 541;
 Best Local Similarity 37.58; Pred. No. 7;
 Matches 9; Conservative 5; Mismatches 3; Indels 7; Gaps 1;

OY 9 DDKASDLIAQIVRRACSDRRRCRW 32
 DB 461 DNKTFEI-----SCCSDRHCKW 477

RESULT 9

| | |
|----|---|
| DE | DNA METHYLASE. |
| GN | MLL9333. |
| OS | Rhizobium loti (Mesorhizobium loti). |
| OC | Plasmid pMLA. |
| QC | Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group; |
| OC | Phyllobacteriaceae; Mesorhizobium. |
| OX | NCBI_TaxID=381; |
| RN | [1] |
| RP | SEQUENCE FROM N.A. |
| RC | STRAIN=MAFF303099; |
| RX | MEDLINE=21082930; PubMed=11214968; |
| RA | Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S., |
| RA | Tanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T., |
| RA | Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., |
| RA | Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M., |
| RA | Takeuchi C., Yamada M., Tabata S., |
| RT | "Complete genome structure of the nitrogen-fixing symbiotic bacterium |
| RT | Mesorhizobium loti."; |
| RL | DNA Res. 7:331-338(2000). |
| RL | EMBL; AP003015; BAB54701.1; - |
| DR | InterPro: IPR002296; N12N6_mtfase. |
| DR | InterPro: IPR000051; SAM_bind. |
| DR | PRINTS; PR00507; N12N6MTFRASE. |
| KW | Methyltransferase; Plasmid; Complete proteome. |
| SQ | SEQUENCE 1011 AA; 111099 MW; 88F2546AC91CEBE7 CRC64; |

| | | | | |
|-----------------------|--------------|---------------|---------------|-------------------|
| Query Match | 27.7%; | Score 54.5; | DB 16; | Length 1011; |
| Best Local Similarity | 50.0%; | Pred. No. 26; | | |
| Matches 13; | Conservative | 2; | Mismatches 8; | Indels 3; Gaps 1; |

| | |
|----|-----------------------------------|
| QY | 8 ADDKASDLIAQIVRRACCSDRRCRW 33 |
| DB | 947 APDKGRDPGAE---RICSRDRCRYR 969 |
| | : : : |

| |
|--|
| RESULT 12 |
| O96440 |
| ID O96440 PRELIMINARY; PRT; 294 AA. |
| AC O96440; |
| DT 01-MAY-1999 (TEMBLrel. 10, Created) |
| DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update) |
| DT 01-MAY-1999 (TEMBLrel. 10, Last annotation update) |
| DE MAJOR SURFACE PROTEIN ASSOCIATED PROTEIN. |
| GN MAG. |
| OS Leishmania chagasi. |
| OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania. |
| ON NCBI_TaxID=5669; |
| OX [1] |
| RP SEQUENCE FROM N.A. |
| RX MEDLINE=99020258; PubMed=9803417; |
| RA McCoy J.J., Beetham J.K., Ochs D.E., Donelson J.E., Wilson M.E.; |
| RT "Regulatory sequences and a novel gene in the msp (GP63) gene cluster |
| RT of Leishmania chagasi." |
| RL Mol. Biochem. Parasitol. 95:251-265(1998). |
| RL EMBL; AF058760; AAD05273.1; - |
| SQ SEQUENCE 294 AA; 31958 MW; 00ADD611D2B93BCE CRC64; |

| | | | | |
|-----------------------|--------------|----------------|----------------|-------------------|
| Query Match | 27.4%; | Score 54; | DB 5; | Length 294; |
| Best Local Similarity | 35.5%; | Pred. No. 8.5; | | |
| Matches 11; | Conservative | 4; | Mismatches 10; | Indels 6; Gaps 1; |

| | |
|----|--|
| QY | 2 DGRNAPADDRKASDLIAQIVRRACCSDRRCRW 32 |
| | : : : |
| DB | 269 EGRTAGAPTATQA-----KTRCCARWRCRW 293 |

| |
|--|
| RESULT 13 |
| O9LA88 |
| ID O9LA88 PRELIMINARY; PRT; 314 AA. |
| AC O9LA88; |
| DT 01-OCT-2000 (TEMBLrel. 15, Created) |

| | | |
|----|---|--|
| AC | Q9RDP5; | |
| AD | 01-MAY-2000 (TReMBLrel. 13, Created) | |
| AE | 01-MAY-2000 (TReMBLrel. 13, Last sequence update) | |
| AF | 01-OCT-2001 (TReMBLrel. 18, Last annotation update) | |
| AG | PUTATIVE LIPOPROTEIN. | |
| AH | SC4A7.20. | |
| AI | Streptomyces coelicolor. | |
| AJ | Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; | |
| AK | Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces | |
| AL | NCBI_TaxID=1902; | |
| AM | [1] | |
| AN | SEQUENCE FROM N.A. | |
| AO | STRAIN=A3(2); | |
| AP | Seeger K.J., Harris D.; | |
| AQ | Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases. | |
| AR | [2] | |
| AS | SEQUENCE FROM N.A. | |
| AT | STRAIN=A3(2); | |
| AV | Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.; | |
| AW | Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases. | |
| AX | [3] | |
| AY | SEQUENCE FROM N.A. | |
| AZ | STRAIN=A3(2); | |
| BA | MEDLINE=97000351; PubMed=8843436; | |
| BB | Redenbach M., Kleser H.M., Denapaire D., Eichner A., Cullum J., | |
| BC | Kinashi H., Hopwood D.A.; | |
| BD | *A set of ordered cosmids and a detailed genetic and physical map | |
| BE | of the 8 Mb Streptomyces coelicolor A3(2) chromosome.*; | |
| BF | Mol. Microbiol. 21:77-96(1996). | |
| BG | EMBL; AL133423; CAB62724.1; | |
| BH | InterPro; IPR001087; Lipase_GDSL. | |
| BI | Prfam; PF00657; Lipase_GDSL; 1. | |
| BJ | Lipoprotein. | |
| BK | SEQUENCE 305 AA; 32696 MW; 387D92F56G388F9C CRC64; | |
| BL | | |
| BM | | |
| BN | | |
| BO | | |
| BP | | |
| BQ | | |
| BR | | |
| BS | | |
| BT | | |
| BU | | |
| BV | | |
| BW | | |
| BX | | |
| BY | | |
| BZ | | |
| CA | | |
| CB | | |
| CC | | |
| CD | | |
| CE | | |
| CF | | |
| CG | | |
| CH | | |
| CI | | |
| CJ | | |
| CK | | |
| CL | | |
| CM | | |
| CN | | |
| CO | | |
| CP | | |
| CQ | | |
| CR | | |
| CS | | |
| CT | | |
| CU | | |
| CV | | |
| CW | | |
| CX | | |
| CY | | |
| CA | | |
| CB | | |
| CC | | |
| CD | | |
| CE | | |
| CF | | |
| CG | | |
| CH | | |
| CI | | |
| CJ | | |
| CK | | |
| CL | | |
| CM | | |
| CN | | |
| CO | | |
| CP | | |
| CQ | | |
| CR | | |
| CS | | |
| CT | | |
| CU | | |
| CV | | |
| CW | | |
| CX | | |
| CY | | |
| CA | | |
| CB | | |
| CC | | |
| CD | | |
| CE | | |
| CF | | |
| CG | | |
| CH | | |
| CI | | |
| CJ | | |
| CK | | |
| CL | | |
| CM | | |
| CN | | |
| CO | | |
| CP | | |
| CQ | | |
| CR | | |
| CS | | |
| CT | | |
| CU | | |
| CV | | |
| CW | | |
| CX | | |
| CY | | |
| CA | | |
| CB | | |
| CC | | |
| CD | | |
| CE | | |
| CF | | |
| CG | | |
| CH | | |
| CI | | |
| CJ | | |
| CK | | |
| CL | | |
| CM | | |
| CN | | |
| CO | | |
| CP | | |
| CQ | | |
| CR | | |
| CS | | |
| CT | | |
| CU | | |
| CV | | |
| CW | | |
| CX | | |
| CY | | |
| CA | | |
| CB | | |
| CC | | |
| CD | | |
| CE | | |
| CF | | |
| CG | | |
| CH | | |
| CI | | |
| CJ | | |
| CK | | |
| CL | | |
| CM | | |
| CN | | |
| CO | | |
| CP | | |
| CQ | | |
| CR | | |
| CS | | |
| CT | | |
| CU | | |
| CV | | |
| CW | | |
| CX | | |
| CY | | |
| CA | | |
| CB | | |
| CC | | |
| CD | | |
| CE | | |
| CF | | |
| CG | | |
| CH | | |
| CI | | |
| CJ | | |
| CK | | |
| CL | | |
| CM | | |
| CN | | |
| CO | | |
| CP | | |
| CQ | | |
| CR | | |
| CS | | |
| CT | | |
| CU | | |
| CV | | |
| CW | | |
| CX | | |
| CY | | |
| CA | | |
| CB | | |

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 12:20:53 ; Search time 51.6 Seconds
(without alignments)
25.831 Million cell updates/sec

Title: US-09-493-795A-5
Sequence: 1 ACCSDRRRCXRC 12

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues
tal number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*

- 1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 71 | 100.0 | 12 | 21 | AA1981373 |
| 2 | 71 | 100.0 | 35 | 21 | AA1981554 |
| 3 | 60 | 84.5 | 32 | 21 | AA1981558 |
| 4 | 54 | 76.1 | 12 | 16 | AA1981572 |
| 5 | 54 | 76.1 | 12 | 18 | AA1981579 |
| 6 | 54 | 76.1 | 12 | 18 | AA1981587 |
| 7 | 54 | 76.1 | 12 | 18 | AA1981594 |
| 8 | 54 | 76.1 | 12 | 19 | AA1981592 |
| 9 | 54 | 76.1 | 12 | 20 | AA1981615 |
| 10 | 54 | 76.1 | 12 | 22 | AA1982215 |
| 11 | 54 | 76.1 | 35 | 21 | AA1981559 |

| | | | | | | |
|----|------|------|-----|----|------------|--------------------|
| 12 | 53 | 74.6 | 32 | 21 | AA1981561 | Cone snail alpha-c |
| 13 | 52 | 73.2 | 13 | 21 | AA1981377 | Cone snail alpha-c |
| 14 | 49 | 69.0 | 14 | 21 | AA1981372 | Cone snail alpha-c |
| 15 | 49 | 69.0 | 35 | 21 | AA1981553 | Cone snail alpha-c |
| 16 | 48 | 67.6 | 48 | 15 | AA1981608 | Agelenopsis peptid |
| 17 | 47 | 66.2 | 19 | 22 | AA1982223 | Toxin peptide SEQ |
| 18 | 46 | 64.8 | 13 | 21 | AA1981378 | Cone snail alpha-c |
| 19 | 46 | 64.8 | 39 | 21 | AA1981601 | Cone snail alpha-c |
| 20 | 46 | 64.8 | 70 | 22 | AA1980640 | Propionibacterium |
| 21 | 46 | 64.8 | 77 | 21 | AA1980808 | Human secreted pro |
| 22 | 46 | 64.8 | 133 | 21 | AA19827973 | Human secreted pro |
| 23 | 46 | 64.8 | 133 | 21 | AA19827974 | Human secreted pro |
| 24 | 46 | 64.8 | 133 | 22 | AA19801087 | Gene 9 Human secre |
| 25 | 46 | 64.8 | 147 | 22 | AA19815890 | Human nervous syst |
| 26 | 46 | 64.8 | 148 | 22 | AA19812580 | Novel human diagno |
| 27 | 46 | 64.8 | 151 | 22 | AA19823434 | Human mdt protein |
| 28 | 46 | 64.8 | 803 | 22 | AA19820106 | Novel human diagno |
| 29 | 45 | 63.4 | 285 | 22 | AA19820056 | Novel human diagno |
| 30 | 45 | 63.4 | 583 | 19 | AA19852118 | Trypanosoma lympho |
| 31 | 44 | 62.0 | 106 | 22 | AA19852378 | Propionibacterium |
| 32 | 44 | 62.0 | 208 | 22 | AA1980932 | Drosophila melanog |
| 33 | 43 | 60.6 | 39 | 21 | AA19821626 | Cone snail alpha-c |
| 34 | 43 | 60.6 | 52 | 21 | AA19845580 | Arabidopsis thalia |
| 35 | 43 | 60.6 | 64 | 22 | AA19829743 | Peptide #2394 enco |
| 36 | 43 | 60.6 | 64 | 22 | AA19834914 | Peptide #2420 enco |
| 37 | 43 | 60.6 | 64 | 22 | AA19820328 | Protein #2327 enco |
| 38 | 43 | 60.6 | 64 | 22 | AA19855725 | Human brain expres |
| 39 | 43 | 60.6 | 64 | 22 | AA19868103 | Human bone marrow |
| 40 | 43 | 60.6 | 64 | 22 | AA19815927 | Peptide #2361 enco |
| 41 | 43 | 60.6 | 64 | 22 | AA19828430 | Peptide #2467 enco |
| 42 | 43 | 60.6 | 64 | 22 | AA19803663 | Peptide #2345 enco |
| 43 | 43 | 60.6 | 101 | 22 | AA19845780 | Propionibacterium |
| 44 | 43 | 60.6 | 130 | 22 | AA1985645 | Human immune/haema |
| 45 | 41.5 | 58.5 | 18 | 22 | AA19892217 | Toxin peptide SEQ |

ALIGNMENTS

RESULT 1
AA1981373
ID AA1981373 standard; Peptide: 12 AA.
XX AC AA1981373;
XX DT 22-JAN-2001 (first entry)
XX DE Cone snail alpha-conotoxin SEQ ID NO: 5.
XX KW Cone snail; alpha-conotoxin; venom; disulphide bond; mood disorder;
XX KW neuronal nicotinic acetylcholine receptor; cardiovascular disorder;
XX KW gastric motility disorder; urinary incontinence; nicotine addiction;
XX KW small cell lung carcinoma.
XX OS Conus imperialis.
XX FH Key Location/Qualifiers
FT Misc-difference 10
FT /label= Trp, OTHER
FT /note= "Trp may be in the D- or L-form, or
FT halo-Trp"
XX WO200044776-A1.
XX PD 03-AUG-2000.
XX PF 28-JAN-2000; 2000WO-US01979.
XX PR 29-JAN-1999; 99US-0118381.
XX PA (UTAH) UNIV UTAH RES FOUND.
XX PA (COGN-) COGNETIX INC.

PI Watkins M, Olivera BM, Hillyard DR, McIntosh JM, Jones RM;

XX WPI; 2000-505965/45.

XX alpha-conotoxin polypeptides derived from the venom of cone snails
PT useful e.g. as neuromuscular blocking agents for use in surgery and for
PT treating unipolar depression -

XX Claim 2; Page 69; 229pp; English.

XX The present invention relates to a number of alpha-conotoxin peptides and
CC their coding sequences from a number of different species of cone snail.
CC These peptides are found in minute quantities in the venom of the snails,
CC and are targeted at the neuronal nicotinic acetylcholine receptors of the
CC nervous system. They usually contain two disulphide bonds, which give
CC them defined conformations, a rarity in molecules this small. The
CC alpha-conotoxins can be used as neuromuscular blocking agents in surgery,
CC and for treating disorders regulated at the neuronal nicotinic
CC acetylcholine receptors, including cardiovascular disorders, gastric
CC motility disorders, urinary incontinence, nicotine addiction, mood
CC disorders such as bipolar disorder, unipolar depression, dysthymia and
CC seasonal affective disorder, and small cell lung carcinoma.

XX Sequence 12 AA;

Query Match 100.0%; Score 71; DB 21; Length 12;

Best Local Similarity 100.0%; Pred. No. 0.0094;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCSDRCRCRXC 12

DB 1 accsdrrcrxc 12

RESULT 2

AAB21554

ID AAB21554 standard; Peptide; 35 AA.

AC AAB21554;

XX 19-JAN-2001 (first entry)

DE Cone snail alpha-conotoxin SEQ ID NO: 236.

XX Cone snail; alpha-conotoxin; venom; disulphide bond; mood disorder;
XX neuronal nicotinic acetylcholine receptor; cardiovascular disorder;
KW gastric motility disorder; urinary incontinence; nicotine addiction;
KW small cell lung carcinoma.

OS Conus imperialis.

XX WO200044776-A1.

XX 03-AUG-2000.

XX 28-JAN-2000; 2000WO-US01979.

XX 29-JAN-1999; 99US-0118381.

XX (UTAH) UNIV UTAH RES FOUND.

PA (COGN-) COGNETIX INC.

XX Watkins M, Olivera BM, Hillyard DR, McIntosh JM, Jones RM;

XX WPI; 2000-505965/45.

DR N-PSDB; AAA89450.

XX alpha-conotoxin polypeptides derived from the venom of cone snails
PT useful e.g. as neuromuscular blocking agents for use in surgery and for
PT treating unipolar depression -

XX Claim 39; Page 46; 229pp; English.

XX The present invention relates to a number of alpha-conotoxin peptides and
CC their coding sequences from a number of different species of cone snail.
CC These peptides are found in minute quantities in the venom of the snails,
CC and are targeted at the neuronal nicotinic acetylcholine receptors of the
CC nervous system. They usually contain two disulphide bonds, which give
CC them defined conformations, a rarity in molecules this small. The
CC alpha-conotoxins can be used as neuromuscular blocking agents in surgery,
CC and for treating disorders regulated at the neuronal nicotinic
CC acetylcholine receptors, including cardiovascular disorders, gastric
CC motility disorders, urinary incontinence, nicotine addiction, mood
CC disorders such as bipolar disorder, unipolar depression, dysthymia and
CC seasonal affective disorder, and small cell lung carcinoma.

XX Sequence 35 AA;

Query Match 100.0%; Score 71; DB 21; Length 35;

Best Local Similarity 91.7%; Pred. No. 0.022;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCSDRCRCRXC 12

DB 23 accsdrrcrwrc 34

RESULT 3

AAB21558

ID AAB21558 standard; Peptide; 32 AA.

XX AAB21558;

XX 19-JAN-2001 (first entry)

DE Cone snail alpha-conotoxin SEQ ID NO: 244.

XX Cone snail; alpha-conotoxin; venom; disulphide bond; mood disorder;
XX neuronal nicotinic acetylcholine receptor; cardiovascular disorder;
KW gastric motility disorder; urinary incontinence; nicotine addiction;
KW small cell lung carcinoma.

XX Conus regius.

XX WO200044776-A1.

XX 03-AUG-2000.

XX 28-JAN-2000; 2000WO-US01979.

XX 29-JAN-1999; 99US-0118381.

XX (UTAH) UNIV UTAH RES FOUND.

PA (COGN-) COGNETIX INC.

XX Watkins M, Olivera BM, Hillyard DR, McIntosh JM, Jones RM;

XX WPI; 2000-505965/45.

DR N-PSDB; AAA89454.

XX alpha-conotoxin polypeptides derived from the venom of cone snails
PT useful e.g. as neuromuscular blocking agents for use in surgery and for
PT treating unipolar depression -

XX Claim 39; Page 47; 229pp; English.

XX The present invention relates to a number of alpha-conotoxin peptides and
CC their coding sequences from a number of different species of cone snail.
CC These peptides are found in minute quantities in the venom of the snails,
CC and are targeted at the neuronal nicotinic acetylcholine receptors of the
CC nervous system. They usually contain two disulphide bonds, which give
CC them defined conformations, a rarity in molecules this small. The
CC alpha-conotoxins can be used as neuromuscular blocking agents in surgery,
CC and for treating disorders regulated at the neuronal nicotinic

CC acetylcholine receptors, including cardiovascular disorders, gastric
 CC motility disorders, urinary incontinence, nicotine addiction, mood
 CC disorders such as bipolar disorder, unipolar depression, dysthymia and
 CC seasonal affective disorder, and small cell lung carcinoma.

SQ Sequence 32 AA;

Query Match 84.5%; Score 60; DB 21; Length 32;
 Best Local Similarity 81.8%; Pred. No. 0.44; Indels 0; Gaps 0;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCSDRRCRXRC 12

|||||:||

Db 21 ccspdrcwrc 31

RESULT 4

AAW75272
 ID AAW75272 standard; peptide; 12 AA.

AC AAW75272;

DT 21-DEC-1995 (first entry)

DE A-lineage conotoxin U002 peptide.

KW Conotoxin; neuromuscular; synapse; signal transmission; inhibitor.

OS Conus imperialis.

FH Key Location/Qualifiers

Modified-site 12

FT /note= "preferably amidated"

XX W09511256-A1.

XX 27-APR-1995.

XX 19-OCT-1994; 94WO-US11927.

XX 19-OCT-1993; 93US-0137800.

XX (UTAH) UNIV UTAH RES FOUND.

PI Cruz LJ, Hillyard DR, McIntosh JM, Olivera BM, Santos AD;

XX WPI; 1995-170189/22.

XX New A-lineage conotoxin peptide(s) - which inhibit synaptic
 transmission at the neuromuscular junction or are active against
 potassium or sodium channels

PS Claim 1; Page 40; 66pp; English.

XX The kappa-conotoxin, alpha conotoxin and alpha-like conotoxin
 CC peptides all belong to a group of peptides known as the A-lineage
 CC conotoxin peptides. The A-lineage conotoxin peptides have a wide
 CC variety of pharmacological uses. The A-lineage conotoxin peptides
 CC claimed (AAW75264-R75293) are useful for the inhibition of synaptic
 CC transmission at neuromuscular junctions by blocking nicotinic acetyl
 CC choline receptors and they also have activity against voltage-gated Na
 CC and K channels.

SQ Sequence 12 AA;

Query Match 76.1%; Score 54; DB 16; Length 12;
 Best Local Similarity 72.7%; Pred. No. 1.1; Indels 0; Gaps 0;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCSDRRCRXRC 12

|||||:||

Db 2 ccspdrcwrc 12

RESULT 5

AAW24879
 ID AAW24879 standard; peptide; 12 AA.

XX AAW24879;

DT 15-OCT-1997 (first entry)

DE Predatory cone snail venom alpha-conotoxin U002.

XX Conotoxin; venom; predatory; cone snail; Conus; A-lineage; inhibitor;
 KW synaptic transmission; neuromuscular junction; block; alpha-conotoxin;
 KW nicotinic acetylcholine receptor; kappa-conotoxin; voltage-sensitive
 KW potassium CHANNEL; sodium channel.

XX Conus imperialis.

FH Key Location/Qualifiers

Modified-site 12

FT /note= "amidated C-terminus"

XX US5633347-A.

XX 27-MAY-1997.

XX 29-JUN-1993; 93US-0084848.

XX 07-JUN-1995; 95US-0480750.

XX 29-JUN-1993; 93US-0084848.

XX 19-OCT-1993; 93US-0137800.

XX (UTAH) UNIV UTAH RES FOUND.

XX Cruz LJ, Hillyard DR, McIntosh JM, Olivera BM, Santos AO;

XX WPI; 1997-309336/28.

XX New kappa-conotoxin peptide(s) - present in venom of fish-hunting
 cone snail

XX Disclosure; Column 4; 37pp; English.

XX The peptides AAW24878-W24900 represent novel toxin peptides isolated
 CC from the venom of various predatory cone snails of the genus Conus. The
 CC peptides are A-lineage conotoxin peptides which fall into 3 groups
 CC dependent on their amino acid sequences: (i) alpha-3/5 have a core
 CC sequence CCXXXCXXXXXC where X is any amino acid; (ii) alpha-4/7 have a
 CC core sequence CCXXXCCXXXXXXC; and (iii) kappa-7/21/3 have the core
 CC sequence CCXXXXXXCCXXCCXXC. The peptide presented here was isolated
 CC from Conus imperialis and falls into a novel alpha-4/3 category.
 CC Alpha-conotoxin peptides are potent inhibitors of synaptic transmission
 CC at the neuromuscular junction by blocking nicotinic acetylcholine
 CC receptors, whereas kappa-conotoxins have activities against
 CC voltage-sensitive potassium or sodium channels.

SQ Sequence 12 AA;

Query Match 76.1%; Score 54; DB 18; Length 12;
 Best Local Similarity 72.7%; Pred. No. 1.1; Indels 0; Gaps 0;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCSDRRCRXRC 12

|||||:||

Db 2 ccspdrcwrc 12

RESULT 6

AAW09447
 ID AAW09447 standard; peptide; 12 AA.

XX AAW09447;
 AC 27-AUG-1997 (first entry)
 DT Alpha-conotoxin peptide U002, targets nicotinic neuronal receptors.
 DE Conotoxin; specificity; nicotinic neuronal receptor; affinity;
 DE targeting; diagnosis; small cell lung carcinoma; SCLC.
 KW Conus imperialis.
 OS
 XX Key Location/Qualifiers
 FH Modified-site 12 /note= "the C-terminus is preferably amidated"
 FT
 FT W09640211-A1.
 XX 19-DEC-1996.
 PN 04-JUN-1996; 96WO-US07962.
 PD 07-JUN-1995; 95US-0487174.
 PR 29-JUN-1993; 93US-0084848.
 PR 19-OCT-1993; 93US-0137800.
 XX (UTAH) UNIV UTAH RES FOUND.
 PA Cruz LJ, Hillyard DR, McIntosh JM, Olivera BM, Santos AS;
 PI WPI; 1997-051898/05.
 XX New use of alpha-conotoxins MII and U002 - for treating and
 DR detecting small cell lung carcinoma
 PT Example 1; Page 6; 29pp; English.
 XX Alpha-conotoxins MII and U002 have a lower affinity for neuromuscular
 CC receptors, than to nicotinic neuronal receptors. This makes them
 CC useful for targeted treatment of small cell lung carcinoma (SCLC), as
 CC cells of this carcinoma express cholinergic nicotinic receptors. The
 CC peptides are administered intravenously or intramuscularly at a
 CC preferred dose of 500 nmoles. When labelled (e.g. with radioactive
 CC iodine), MII and U002 can be used for detecting SCLC tumours.
 XX Sequence 12 AA;
 SQ

Query Match 76.1%; Score 54; DB 18; Length 12;
 Best Local Similarity 72.7%; Pred. No. 1.1;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 2 CCSDRRCXRC 12
 ||||| || :||
 Db 2 ccspcawrc.12

RESULT 7
 AAW12734
 ID AAW12734 standard; Peptide; 12 AA.
 XX AAW12734;
 XX 16-APR-1997 (first entry)
 DT Alpha-conotoxin peptide U002.
 DE Polymerase chain reaction; PCR; primer; amplify; conotoxin; Conus;
 KW inhibitor; synaptic transmission; neuromuscular junction; sodium channel;
 KW nicotinic acetylcholine receptor; potassium channel; muscle relaxant;
 KW myasthenia gravis; small cell lung cancer; therapy.
 XX Conus imperialis.
 OW

XX Key Location/Qualifiers
 FH Modified-site 12 /note= "amidated"
 FT
 XX US5589340-A.
 XX 31-DEC-1996.
 XX 29-JUN-1993; 93US-0084848.
 XX 07-JUN-1995; 95US-0477383.
 PR 29-JUN-1993; 93US-0084848.
 PR 19-OCT-1993; 93US-0137800.
 XX (UTAH) UNIV UTAH RES FOUND.
 PA Cruz LJ, Hillyard DR, McIntosh JM, Olivera BM, Santos AD;
 PI WPI; 1997-076840/07.
 DR Identifying nucleic acid encoding A-lineage conotoxin peptide(s) by
 PT amplification - uses primers corresponding to conserved regions in
 PT the signal sequence and 3'-untranslated regions, useful e.g. in
 PT treatment of small cell lung cancer
 XX Disclosure; Column 4; 36pp; English.
 PS
 XX AAW12726-W12769 represent conotoxin peptides. This sequence represents
 CC the U002 alpha-conotoxin peptide isolated from Conus imperialis. These
 CC sequences are identified using the method of the invention. The method
 CC of the invention is for identifying DNA encoding A-lineage conotoxin
 CC peptides by subjecting Conus nucleic acid to amplification with primer
 CC sequences (see AAF59714 and AAF59715). The primers are specific for the
 CC signal sequence and 3'-untranslated (3'UTR) regions of the conotoxin
 CC gene, which are highly homologous between conotoxins, and are therefore
 CC suitable sites for detection. A-lineage conotoxins include alpha-
 CC conotoxins, and kappa-conotoxins. Alpha-conotoxins are powerful
 CC inhibitors of synaptic transmission at the neuromuscular junction, and
 CC are usually nicotinic acetylcholine receptor blockers. Kappa-conotoxins
 CC act on the voltage sensitive sodium and potassium channels. The
 CC conotoxins identified can be used as muscle relaxants, in the diagnosis
 CC of myasthenia gravis, and for the treatment or diagnosis of small cell
 CC lung cancer. For the treatment of small cell lung cancer, the conotoxin
 CC peptides act by binding to the nicotinic receptors, and thereby blocking
 CC the nicotine/cytosine stimulated release of the mitogen
 CC 5-hydroxytryptamine.
 XX Sequence 12 AA;
 SQ

Query Match 76.1%; Score 54; DB 18; Length 12;
 Best Local Similarity 72.7%; Pred. No. 1.1;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 2 CCSDRRCXRC 12
 ||||| || :||
 Db 2 ccspcawrc 12

RESULT 8
 AAW57902
 ID AAW57902 standard; peptide; 12 AA.
 XX AAW57902;
 AC 25-SEP-1998 (first entry)
 DT Conotoxin peptide ImI.
 DE Conotoxin peptide; ImI; MII; cardiovascular agent; altered heart rate;
 KW altered blood pressure; nicotinic acetylcholine receptor antagonist;
 KW B neurone blocker; venom; marine snail; C neurone blocker;

KW sympathetic impulse.
OS Conus Imperialis.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 2..8
FT Disulfide-bond 3..12
XX
PN W09822126-A1.
XX
PD 28-MAY-1998.
XX
PF 17-NOV-1997; 97WO-US20669.
XX
PR 18-NOV-1996; 96US-0031141.
XX
PA (UTAH) UNIV UTAH RES FOUND.
XX
PI McIntosh JM, Olivera BM, Yoshikami D;
PR WPI; 1998-322346/28.
XX
PT Use of the conotoxin peptide(s) ImI and MII - as agents which can
PT regulate heart rate or blood pressure
XX
PS Claim 1; Page 4; 24pp; English.
XX
CC This sequence represents the conotoxin peptide ImI. This sequence and
CC the MII conotoxin peptide (see AAW57903) can be used in the method of
CC the invention for the treatment of a patient who has an altered heart
CC rate or an altered blood pressure. The peptides are found in the venom of
CC marine snails of the genus Conus. They are active as nicotinic
CC acetylcholine receptor antagonists. They differentially block the B and C
CC neurons, and are thus able to differentially block sympathetic impulses
CC to the heart affecting the heart rate and blood pressure. The above
CC agents are capable of discretely affecting the heart rate or blood
CC pressure, without affecting other muscles.
XX
SQ Sequence 12 AA;

Query Match 76.1%; Score 54; DB 19; Length 12;
Best Local Similarity 72.7%; Pred. No. 1.1;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 CCSDRCRCXRC 12
DB ||||| :||
2 ccsdprcawrc 12

RESULT 9
AAY24165
ID AAY24165 standard; peptide; 12 AA.
XX
AC AAY24165;
XX
DT 10-SEP-1999 (first entry)
XX
DE Alpha-conotoxin peptide SEQ ID NO:13.
XX
KW Alpha-conotoxin; neuronal nicotinic acetylcholine receptor; nAChR;
KW small cell lung carcinoma; cardiovascular disorder; nicotine addiction;
KW gastric motility disorder; urinary incontinence; mood disorder;
KW bipolar disorder; unipolar depression; dysthymia;
KW seasonal effective disorder.
XX
OS Conus Imperialis.
XX
PN W09933482-A1.
XX
PD 08-JUL-1999.
XX
PF 23-DEC-1998; 98WO-US27367.

XX
PR 03-APR-1998; 98US-0080588.
PR 31-DEC-1997; 97US-0070153.
XX
PA (UTAH) UNIV UTAH RES FOUND.
XX
PI Cartier GE, Luo S, McIntosh JM, Olivera BM, Yoshikami D;
XX WPI; 1999-405367/34.
DR
XX
PT Alpha-conotoxin peptides that are used to treat disorders regulated
PT at neuronal nicotinic acetylcholine receptors
XX
PS Claim 28; Page 6; 40pp; English.
XX
CC The present sequence represents a specifically claimed example of an
CC alpha-conotoxin, which can be used to treat disorders regulated at
CC neuronal nicotinic acetylcholine receptors (nAChR). The alpha-conotoxins
CC are useful for preparing a pharmaceutical composition for treating
CC disorders regulated at neuronal nAChR, especially alpha 3 beta 2,
CC alpha 3 beta 4 or alpha 7-containing nAChR. Disorders that can be
CC treated include cardiovascular disorders, a gastric motility disorder,
CC urinary incontinence, nicotine addiction, a mood disorder or small cell
CC lung carcinoma. Mood disorders include bipolar disorder, unipolar
CC depression, dysthymia and seasonal effective disorder. The alpha-
CC conotoxins can also be used for diagnosis of small cell lung carcinoma.
CC The alpha-conotoxin antagonists are able to discriminate between non-
CC symmetrical ligand binding interfaces present on the nAChR. The alpha-
CC conotoxin has the ability to potentially block any receptor containing a
CC alpha beta subunit interface, regardless of what other subunits may be
CC present in the receptor complex.
XX
SQ Sequence 12 AA;

Query Match 76.1%; Score 54; DB 20; Length 12;
Best Local Similarity 72.7%; Pred. No. 1.1;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 CCSDRCRCXRC 12
DB ||||| :||
2 ccsdprcawrc 12

RESULT 10
AAB92215
ID AAB92215 standard; Peptide; 12 AA.
XX
AC AAB92215;
XX
DT 22-JUN-2001 (first entry)
XX
DE Toxin peptide SEQ ID NO:1391.
XX
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW blood component; modification; succinimidy; maleimido group; amino;
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN W0200069900-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-US13576.
XX
PR 17-MAY-1999; 99US-0134406.
PR 10-SEP-1999; 99US-0153406.
PR 15-OCT-1999; 99US-0159783.
XX
PA (CONJ-) CONJUCHEM INC.

PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
 XX WPI; 2001-112059/12.
 XX
 PT Modifying and attaching therapeutic peptides to albumin prevents
 PT peptidase degradation, useful for increasing length of in vivo activity
 XX
 PS Disclosure; Page 651; 733pp; English.
 XX
 CC The present invention describes a modified therapeutic peptide (I)
 CC comprising a therapeutically active amino acid region (III) and a
 CC reactive group (II) (e.g. succinimide and maleimide groups) attached to
 CC a less therapeutically active amino acid region (IV), which covalently
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a
 CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
 CC factors and neurotransmitters, to protect them from peptidase activity
 CC in vivo for the treatment of various disorders. Endogenous therapeutic
 CC peptides are not suitable as drug candidates as they require frequent
 CC administration due to rapid degradation by peptidases in the body.
 CC Modifying and attaching therapeutic peptides to albumin prevents or
 CC reduces the action of peptidases to increase length of activity (half
 CC life) and specificity as bonding to large molecules decreases
 CC intracellular uptake and interference with physiological processes.
 CC AAB90829 to AAB92441 represent peptides which can be used in the
 CC exemplification of the present invention.
 XX
 SQ Sequence 12 AA;

Query Match 76.1%; Score 54; DB 22; Length 12;
 Best Local Similarity 72.7%; Pred. No. 1.1;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CCSDRRCRXRC 12
 |||| || :||
 Db 2 ccspdcawrc 12

RESULT 11

AAB21559
 ID AAB21559 standard; Peptide; 35 AA.

XX
 AC AAB21559;

XX
 DT 19-JAN-2001 (first entry)

XX Cone snail alpha-conotoxin SEQ ID NO: 246.

XX Cone snail; alpha-conotoxin; venom; disulphide bond; mood disorder;
 KW neuronal nicotinic acetylcholine receptor; cardiovascular disorder;
 KW gastric motility disorder; urinary incontinence; nicotine addiction;
 KW small cell lung carcinoma.

XX Conus regius.

XX WO200044776-A1.

XX 03-AUG-2000.

XX 28-JAN-2000; 2000WO-US01979.

XX 29-JAN-1999; 99US-0118381.

XX (UTAH) UNIV UTAH RES FOUND.

XX (COGN-) COGNETIX INC.

XX Watkins M, Olivera BM, Hillyard DR, McIntosh JM, Jones RM;

XX WPI; 2000-505965/45.

XX N-PSDB; AAA89455.

XX

PT alpha-conotoxin polypeptides derived from the venom of cone snails
 PT useful e.g. as neuromuscular blocking agents for use in surgery and for
 PT treating unipolar depression -
 XX
 XX Claim 39; Page 47; 229pp; English.
 XX
 CC The present invention relates to a number of alpha-conotoxin peptides and
 CC their coding sequences from a number of different species of cone snail.
 CC These peptides are found in minute quantities in the venom of the snails,
 CC and are targeted at the neuronal nicotinic acetylcholine receptors of the
 CC nervous system. They usually contain two disulphide bonds, which give
 CC them defined conformations, a rarity in molecules this small. The
 CC alpha-conotoxins can be used as neuromuscular blocking agents in surgery,
 CC and for treating disorders regulated at the neuronal nicotinic
 CC acetylcholine receptors, including cardiovascular disorders, gastric
 CC motility disorders, urinary incontinence, nicotine addiction, mood
 CC disorders such as bipolar disorder, unipolar depression, dysthymia and
 CC seasonal affective disorder, and small cell lung carcinoma.
 XX
 SQ Sequence 35 AA;

Query Match 76.1%; Score 54; DB 21; Length 35;
 Best Local Similarity 72.7%; Pred. No. 2.6;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CCSDRRCRXRC 12

|||| || :||

Db 24 ccspdcawrc 34

RESULT 12

AAB21561
 ID AAB21561 standard; Peptide; 32 AA.

XX
 AC AAB21561;

XX
 DT 19-JAN-2001 (first entry)

XX Cone snail alpha-conotoxin SEQ ID NO: 250.

XX Cone snail; alpha-conotoxin; venom; disulphide bond; mood disorder;
 KW neuronal nicotinic acetylcholine receptor; cardiovascular disorder;
 KW gastric motility disorder; urinary incontinence; nicotine addiction;
 KW small cell lung carcinoma.

XX Conus regius.

XX WO200044776-A1.

XX 03-AUG-2000.

XX 28-JAN-2000; 2000WO-US01979.

XX 29-JAN-1999; 99US-0118381.

XX (UTAH) UNIV UTAH RES FOUND.

XX (COGN-) COGNETIX INC.

XX Watkins M, Olivera BM, Hillyard DR, McIntosh JM, Jones RM;

XX WPI; 2000-505965/45.

XX N-PSDB; AAA89457.

XX alpha-conotoxin polypeptides derived from the venom of cone snails
 PT useful e.g. as neuromuscular blocking agents for use in surgery and for
 PT treating unipolar depression -

XX Claim 39; Page 47; 229pp; English.

XX The present invention relates to a number of alpha-conotoxin peptides and
 CC their coding sequences from a number of different species of cone snail.
 CC These peptides are found in minute quantities in the venom of the snails,

CC and are targeted at the neuronal nicotinic acetylcholine receptors of the
CC nervous system. They usually contain two disulphide bonds, which give
CC them defined conformations, a rarity in molecules this small. The
CC alpha-conotoxins can be used as neuromuscular blocking agents in surgery,
CC and for treating disorders regulated at the neuronal nicotinic
CC acetylcholine receptors, including cardiovascular disorders, gastric
CC motility disorders, urinary incontinence, nicotine addiction, mood
CC disorders such as bipolar disorder, unipolar depression, dysthymia and
CC seasonal affective disorder, and small cell lung carcinoma.

XX Sequence 32 AA;

Query Match 74.6%; Score 53; DB 21; Length 32;
Best Local Similarity 63.6%; Pred. No. 3.2;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 2 CCSDRRCRXRC 12
|||||
21 ccsdprckhqc 31

RESULT 13

AAB21377
ID AAB21377 standard; Peptide; 13 AA.

XX
AC AAB21377;

XX 22-JAN-2001 (first entry)

XX Cone snail alpha-conotoxin SEQ ID NO: 9.

XX Cone snail; alpha-conotoxin; venom; disulphide bond; mood disorder;
KW neuronal nicotinic acetylcholine receptor; cardiovascular disorder;
KW gastric motility disorder; urinary incontinence; nicotine addiction;
KW small cell lung carcinoma.

XX Conus regius.

XX Key Location/Qualifiers

FT Misc-difference 6

FT /label= Pro, OTHER

FT /note= "hydroxy-Pro"

FT Misc-difference 10

FT /label= Tyr, OTHER

FT /note= "nor-Tyr, mono-halo-Tyr, di-halo-Tyr,

FT O-sulpho-Tyr, O-phospho-Tyr or nitro-Tyr"

XX WO200044776-A1.

XX 03-AUG-2000.

XX 28-JAN-2000; 2000WO-US01979.

XX 29-JAN-1999; 99US-0118381.

XX (UTAH) UNIV UTAH RES FOUND.

XX (COGN-) COGNETIX INC.

XX Watkins M, Olivera BM, Hillyard DR, McIntosh JM, Jones RM;

XX WPI; 2000-505965/45.

XX alpha-conotoxin polypeptides derived from the venom of cone snails
PT useful e.g. as neuromuscular blocking agents for use in surgery and for
PT treating unipolar depression -

XX Claim 2; Page 70; 229pp; English.

XX The present invention relates to a number of alpha-conotoxin peptides and
CC their coding sequences from a number of different species of cone snail.
CC These peptides are found in minute quantities in the venom of the snails,
CC and are targeted at the neuronal nicotinic acetylcholine receptors of the

CC nervous system. They usually contain two disulphide bonds, which give
CC them defined conformations, a rarity in molecules this small. The
CC alpha-conotoxins can be used as neuromuscular blocking agents in surgery,
CC and for treating disorders regulated at the neuronal nicotinic
CC acetylcholine receptors, including cardiovascular disorders, gastric
CC motility disorders, urinary incontinence, nicotine addiction, mood
CC disorders such as bipolar disorder, unipolar depression, dysthymia and
CC seasonal affective disorder, and small cell lung carcinoma.

XX Sequence 13 AA;

Query Match 73.2%; Score 52; DB 21; Length 13;
Best Local Similarity 90.9%; Pred. No. 2.1;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 CCSDRRCRXRC 12
|||||
2 ccsdxcrcrc 12

RESULT 14

AAB21372
ID AAB21372 standard; Peptide; 14 AA.

XX
AC AAB21372;

XX 22-JAN-2001 (first entry)

XX Cone snail alpha-conotoxin SEQ ID NO: 4.

XX Cone snail; alpha-conotoxin; venom; disulphide bond; mood disorder;
KW neuronal nicotinic acetylcholine receptor; cardiovascular disorder;
KW gastric motility disorder; urinary incontinence; nicotine addiction;
KW small cell lung carcinoma.

XX Conus imperialis.

XX Key Location/Qualifiers

FT Misc-difference 2

FT /label= Glu, OTHER

FT /note= "gamma-carboxy-Glu"

FT Misc-difference 11

FT /label= Lys, OTHER

FT /note= "N-methyl-Lys, N,N-dimethyl-Lys or

FT N,N,N-trimethyl-Lys"

XX WO200044776-A1.

XX 03-AUG-2000.

XX 28-JAN-2000; 2000WO-US01979.

XX 29-JAN-1999; 99US-0118381.

XX (UTAH) UNIV UTAH RES FOUND.

XX (COGN-) COGNETIX INC.

XX Watkins M, Olivera BM, Hillyard DR, McIntosh JM, Jones RM;

XX WPI; 2000-505965/45.

XX alpha-conotoxin polypeptides derived from the venom of cone snails
PT useful e.g. as neuromuscular blocking agents for use in surgery and for
PT treating unipolar depression -

XX Claim 2; Page 69; 229pp; English.

XX The present invention relates to a number of alpha-conotoxin peptides and
CC their coding sequences from a number of different species of cone snail.
CC These peptides are found in minute quantities in the venom of the snails,
CC and are targeted at the neuronal nicotinic acetylcholine receptors of the
CC nervous system. They usually contain two disulphide bonds, which give

CC them defined conformations, a rarity in molecules this small. The
CC alpha-conotoxins can be used as neuromuscular blocking agents in surgery,
CC and for treating disorders regulated at the neuronal nicotinic
CC acetylcholine receptors, including cardiovascular disorders, gastric
CC motility disorders, urinary incontinence, nicotine addiction, mood
CC disorders such as bipolar disorder, unipolar depression, dysthymia and
CC seasonal affective disorder, and small cell lung carcinoma.
XX
SQ Sequence 14 AA;

Query Match 69.0%; Score 49; DB 21; Length 14;
Best Local Similarity 72.7%; Pred. No. 5.1;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCSDRRCRXRC 12
Db 3 ccdsrccgnc 13

SUBT 15

ID AAB21553 standard; Peptide; 35 AA.

XX
AC/

XX
AAB21553;

DT 19-JAN-2001 (first entry)

DE Cone snail alpha-conotoxin SEQ ID NO: 234.

XX
KW Cone snail; alpha-conotoxin; venom; disulphide bond; mood disorder;
KW neuronal nicotinic acetylcholine receptor; cardiovascular disorder;
KW gastric motility disorder; urinary incontinence; nicotine addiction;
KW small cell lung carcinoma.

XX
OS Conus imperialis.

XX
PN W0200044776-A1.

XX
PD 03-AUG-2000.

XX
PF 28-JAN-2000; 2000WO-US01979.

XX
PR 29-JAN-1999; 99US-0118381.

XX
PA (UTAH) UNIV UTAH RES FOUND.

XX
PA (COGN-) COGNETIX INC.

Watkins M, Olivera BM, Hillyard DR, McIntosh JM, Jones RM;

WPI: 2000-505965/45.

DR N-PSDB; AA89449.

XX
PT alpha-conotoxin polypeptides derived from the venom of cone snails
PT useful e.g. as neuromuscular blocking agents for use in surgery and for
PT treating unipolar depression .
XX
PS Claim 39; Page 46; 229pp; English.

XX
CC The present invention relates to a number of alpha-conotoxin peptides and
CC their coding sequences from a number of different species of cone snail.
CC These peptides are found in minute quantities in the venom of the snails,
CC and are targeted at the neuronal nicotinic acetylcholine receptors of the
CC nervous system. They usually contain two disulphide bonds, which give
CC them defined conformations, a rarity in molecules this small. The
CC alpha-conotoxins can be used as neuromuscular blocking agents in surgery,
CC and for treating disorders regulated at the neuronal nicotinic
CC acetylcholine receptors, including cardiovascular disorders, gastric
CC motility disorders, urinary incontinence, nicotine addiction, mood
CC disorders such as bipolar disorder, unipolar depression, dysthymia and
CC seasonal affective disorder, and small cell lung carcinoma.

XX
SQ Sequence 35 AA;

Query Match 69.0%; Score 49; DB 21; Length 35;
Best Local Similarity 63.6%; Pred. No. 10;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCSDRRCRXRC 12
Db 24 ccdsarccgnc 34

Search completed: July 1, 2002, 12:24:52
Job time: 239 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 12:22:43 ; Search time 20.57 Seconds
(without alignments)
14,249 Million cell updates/sec

Title: US-09-493-795A-5
Perfect score: 71
Sequence: 1 ACCSDRRRCXRRC 12

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

tal number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Match | Query Length | DB ID | Description |
|------------|-------|---------|--------------|-------|-------------------|
| 1 | 54 | 76.1 | 12 | 1 | US-08-137-800-12 |
| 2 | 54 | 76.1 | 12 | 1 | US-08-477-383-12 |
| 3 | 54 | 76.1 | 12 | 1 | US-08-487-174-12 |
| 4 | 54 | 76.1 | 12 | 1 | US-08-480-750-12 |
| 5 | 54 | 76.1 | 12 | 4 | US-09-219-446B-13 |
| 6 | 54 | 76.1 | 12 | 5 | PCT-US96-07962-9 |
| 7 | 48 | 67.6 | 48 | 1 | US-08-379-550-1 |
| 8 | 45 | 63.4 | 31 | 2 | US-08-686-599A-9 |
| 9 | 41 | 57.7 | 19 | 4 | US-09-136-769A-3 |
| 10 | 41 | 57.7 | 19 | 4 | US-09-136-769A-14 |
| 11 | 41 | 57.7 | 29 | 1 | US-08-775-476-1 |
| 12 | 41 | 57.7 | 29 | 1 | US-09-018-799-1 |
| 13 | 41 | 57.7 | 29 | 2 | US-09-099-974-1 |
| 14 | 41 | 57.7 | 58 | 6 | 5189019-8 |
| 15 | 41 | 57.7 | 75 | 6 | 5189019-4 |
| 16 | 41 | 57.7 | 136 | 6 | 5189019-6 |
| 17 | 40.5 | 57.0 | 175 | 1 | US-08-194-180-3 |
| 18 | 40 | 56.3 | 415 | 4 | US-09-006-353A-6 |
| 19 | 39 | 54.9 | 46 | 4 | US-08-632-511A-5 |
| 20 | 39 | 54.9 | 46 | 4 | US-09-091-590A-11 |
| 21 | 39 | 54.9 | 46 | 4 | US-09-488-200-5 |
| 22 | 39 | 54.9 | 74 | 1 | US-08-543-238-5 |
| 23 | 39 | 54.9 | 74 | 1 | US-08-420-526-5 |
| 24 | 39 | 54.9 | 202 | 1 | US-08-048-700-2 |
| 25 | 39 | 54.9 | 202 | 1 | US-08-855-261A-3 |
| 26 | 39 | 54.9 | 202 | 3 | US-08-839-711-3 |
| 27 | 39 | 54.9 | 202 | 4 | US-09-227-224-3 |

28 39 54.9 488 1 US-08-243-542-1 Sequence 1, Appli
29 39 54.9 488 1 US-08-477-407-1 Sequence 1, Appli
30 39 54.9 488 1 US-08-484-355-1 Sequence 1, Appli
31 39 54.9 524 1 US-08-243-542-2 Sequence 2, Appli
32 39 54.9 524 1 US-08-477-407-2 Sequence 2, Appli
33 39 54.9 524 1 US-08-484-355-2 Sequence 2, Appli
34 39 54.9 670 1 US-08-243-542-3 Sequence 3, Appli
35 39 54.9 670 1 US-08-477-407-3 Sequence 3, Appli
36 39 54.9 670 1 US-08-484-355-3 Sequence 3, Appli
37 39 54.9 749 2 US-08-568-459A-6 Sequence 6, Appli
38 39 54.9 749 2 US-08-487-826B-6 Sequence 6, Appli
39 39 54.9 769 1 US-08-243-542-4 Sequence 4, Appli
40 39 54.9 769 1 US-08-477-407-4 Sequence 4, Appli
41 39 54.9 769 1 US-08-484-355-4 Sequence 4, Appli
42 39 54.9 1070 4 US-08-697-954-2 Sequence 2, Appli
43 38 53.5 17 4 US-09-604-864-1 Sequence 1, Appli
44 38 53.5 55 2 US-08-369-829A-1 Sequence 1, Appli
45 38 53.5 55 2 US-08-369-829A-17 Sequence 17, Appli

ALIGNMENTS

RESULT 1
US-08-137-800-12
; Sequence 12, Application US/08137800
; Patent No. 5514774
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Hillyard, David R.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Santos, Ameurina D.
; TITLE OF INVENTION: Conotoxin Peptides
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue N.W., Suite 1000
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/137,800
; FILING DATE: 19-OCT-1993
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Innen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24260-104763
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Conus imperialis
; US-08-137-800-12

Query Match 76.1%; Score 54; DB 1; Length 12;
Best Local Similarity 72.7%; Pred. No. 0.5;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCSDRRCXRC 12
Db 2 CCSDPRCAWRC 12

RESULT 2

US-08-477-383-12
; Sequence 12, Application US/08477383
; Patent No. 5589340
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Hillyard, David R.
; APPLICANT: Macintosh, J. Michael
; APPLICANT: Santos, Ameurfino S.
; TITLE OF INVENTION: Conotoxin Peptides
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,383
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/137,800
; FILING DATE: 19-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/084,848
; FILING DATE: 29-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24260-107673
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Conus imperialis
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 12
; OTHER INFORMATION: /note= "The C-terminus is preferably amidated."
US-08-477-383-12

Query Match 76.1%; Score 54; DB 1; Length 12;
Best Local Similarity 72.7%; Pred. NO. 0.5;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCSDRRCXRC 12
Db 2 CCSDPRCAWRC 12

RESULT 4

US-08-480-750-12
; Sequence 12, Application US/08480750
; Patent No. 5633347

RESULT 3

US-08-487-174-12
; Sequence 12, Application US/08487174
; Patent No. 5595972
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Hillyard, David R.
; APPLICANT: Macintosh, J. Michael
; APPLICANT: Santos, Ameurfino S.
; TITLE OF INVENTION: Conotoxin Peptides
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,174
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/137,800
; FILING DATE: 19-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/084,848
; FILING DATE: 29-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24260-107673
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Conus imperialis
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 12
; OTHER INFORMATION: /note= "The C-terminus is preferably amidated."
US-08-487-174-12

Query Match 76.1%; Score 54; DB 1; Length 12;
Best Local Similarity 72.7%; Pred. NO. 0.5;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCSDRRCXRC 12
Db 2 CCSDPRCAWRC 12

GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Hillyard, David R.
; APPLICANT: Macintosh, J. Michael
; APPLICANT: Santos, Ameurfinio S.
; TITLE OF INVENTION: Conotoxin Peptides
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,750
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/137,800
; FILING DATE: 19-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/084,848
; FILING DATE: 29-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24260-107673
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Conus imperialis
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 12
; OTHER INFORMATION: /note= "The C-terminus is
; OTHER INFORMATION: preferably amidated."
US-08-480-750-12

Query Match 76.1%; Score 54; DB 1; Length 12;
Best Local Similarity 72.7%; Pred. No. 0.5;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCSDRRRCXRC 12
| | | | | : | |
Db 2 CCSDPRCAWRC 12

RESULT 5
US-09-219-446B-13
; Sequence 13, Application US/09219446B
; Patent No. 6265541
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Yoshikami, Doju
; APPLICANT: Cartier, G. Edward

; APPLICANT: Luo, Sigin
; APPLICANT: University of Utah Research Foundation
; TITLE OF INVENTION: Uses of Alpha-Conotoxin Peptides
; FILE REFERENCE: Uses of Alpha-Conotoxins
; CURRENT APPLICATION NUMBER: US/09/219,446B
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/080,588
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: US 60/070,153
; PRIOR FILING DATE: 1997-12-31
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 12
; TYPE: PPT
; ORGANISM: Conus imperialis
US-09-219-446B-13

Query Match 76.1%; Score 54; DB 4; Length 12;
Best Local Similarity 72.7%; Pred. No. 0.5;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCSDRRRCXRC 12
| | | | | : | |
Db 2 CCSDPRCAWRC 12

RESULT 6
PCT-US96-07962-9
; Sequence 9, Application PC/TUS9607962
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; TITLE OF INVENTION: Use of Conotoxin Peptides U002 and M11
; TITLE OF INVENTION: for Treating or Detecting Small-Cell Lung Carcinoma
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-WINDOWS
; SOFTWARE: Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/07962
; FILING DATE: 04-JUN-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,174
; FILING DATE: 07-JUN-1995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Conus imperialis
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 12
; OTHER INFORMATION: /note= "The C-terminus is
; OTHER INFORMATION: preferably amidated."

PCT-US96-07962-9

Query Match 76.1%; Score 54; DB 5; Length 12;
Best Local Similarity 72.7%; Pred. No. 0.5;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCSDRRCRXRC 12
||||| :|||
Db 2 CCSDPRCAWRC 12

RESULT 7
US-08-379-550-1
; Sequence 1, Application US/08379550
; Patent No. 5595559
; GENERAL INFORMATION:
; APPLICANT: Phillips, Douglas
; APPLICANT: Kelly, Mary E.
; APPLICANT: Saccomano, Nicholas A.
; APPLICANT: Volkman, Robert A.
; TITLE OF INVENTION: Calcium Channel Blocking Polypeptide
; TITLE OF INVENTION: From Agelenopsis aperta
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pfizer Inc Patent Department
; STREET: Eastern Point Road
; CITY: Groton
; STATE: CT
; COUNTRY: USA
; ZIP: 06340
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379,550
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07919,538
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: McFarlin, D. Stuart
; REGISTRATION NUMBER: 33,736
; REFERENCE/DOCKET NUMBER: PC8172DSM
; TELEPHONE: (203) 441-4905
; TELEFAX: (203) 441-5221
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 48 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Agelenopsis aperta
; TISSUE TYPE: venom
US-08-379-550-1

Query Match 67.6%; Score 48; DB 1; Length 48;
Best Local Similarity 77.8%; Pred. No. 7.1;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCSDRRCRX 10
||| :|||
Db 15 CCHDRRCRC 23

RESULT 8
US-08-686-599A-9
; Sequence 9, Application US/08686599A
; Patent No. 5891439
; GENERAL INFORMATION:
; APPLICANT: Olsson, Thomas
; APPLICANT: Vaidya, Tushar
; APPLICANT: Bakht, Abdel-Moiz
; APPLICANT: Kristensson, Krister
; APPLICANT: Donelson, John E.
; TITLE OF INVENTION: LYMPHOCYTE STIMULATING FACTOR
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22314-3187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/686,599A
; FILING DATE: 26-JUL-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Duffett, Benton S.
; REGISTRATION NUMBER: 22,030
; REFERENCE/DOCKET NUMBER: 003300-383
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-686-599A-9

Query Match 63.4%; Score 45; DB 2; Length 31;
Best Local Similarity 60.0%; Pred. No. 11;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCSDRRCRXR 11
||| :|||
Db 21 CCHDRRCETR 30

RESULT 9
US-09-136-769A-3
; Sequence 3, Application US/09136769A
; Patent No. 6307014
; GENERAL INFORMATION:
; APPLICANT: Furie, Bruce
; APPLICANT: Furie, Barbara
; APPLICANT: Stenflo, Johan
; APPLICANT: Rigby, Alan C.
; APPLICANT: Roepstoft, Peter
; TITLE OF INVENTION: CONOPEPTIDES
; FILE REFERENCE: 50065/002001
; CURRENT APPLICATION NUMBER: US/09/136,769A
; CURRENT FILING DATE: 1998-08-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Conus textile
; FEATURE:
; NAME/KEY: VARIANT

; LOCATION: (2)...(16)
; OTHER INFORMATION: Pro at position 2 is 4Hyp and each Xaa is
; OTHER INFORMATION: Independently selected from Glu and
; OTHER INFORMATION: gamma-carboxyglutamic acid, provided that at least
; OTHER INFORMATION: one Xaa is gamma-carboxyglutamic acid.
; NAME/KEY: VARIANT
; LOCATION: (1)...(19)
; OTHER INFORMATION: Xaa = Any Amino Acid
; NAME/KEY: VARIANT
; LOCATION: (1)...(19)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-136-769A-3

Query Match 57.7%; Score 41; DB 4; Length 19;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CCSDRRCRX 10
||||| :
Db 4 CCSDPRCNS 12

RESULT 10
US-09-136-769A-14
; Sequence 14, Application US/09136769A
; Patent No. 6307014
; GENERAL INFORMATION:
; APPLICANT: Furie, Bruce
; APPLICANT: Furie, Barbara
; APPLICANT: Stenflo, Johan
; APPLICANT: Rigby, Alan C.
; APPLICANT: Roepstoft, Peter
; TITLE OF INVENTION: CONOPEPTIDES
; FILE REFERENCE: 50065/002001
; CURRENT APPLICATION NUMBER: US/09/136.769A
; CURRENT FILING DATE: 1998-08-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Conus textile
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (2)...(16)
; OTHER INFORMATION: Pro at position 2 is 4Hyp and Xaa at positions 3
; OTHER INFORMATION: and 16 is gamma-carboxyglutamic acid.
; NAME/KEY: VARIANT
; LOCATION: (1)...(19)
; OTHER INFORMATION: Xaa = Any Amino Acid
; NAME/KEY: VARIANT
; LOCATION: (1)...(19)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-136-769A-14

Query Match 57.7%; Score 41; DB 4; Length 19;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CCSDRRCRX 10
||||| :
Db 4 CCSDPRCNS 12

RESULT 11
US-08-775-476-1
; Sequence 1, Application US/08775476
; Patent No. 5776896
; GENERAL INFORMATION:
; APPLICANT: Lampe, Richard A.
; TITLE OF INVENTION: Analgesic Peptides from Venom of

; TITLE OF INVENTION: Grammostola Spatulata and Use Thereof
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZENECA Pharmaceuticals
; STREET: 1800 Concord Pike
; CITY: Wilmington
; STATE: DE
; COUNTRY: United States
; ZIP: 19850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/775,476
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hohenschutz, Liza D.
; REGISTRATION NUMBER: 33,712
; REFERENCE/DOCKET NUMBER: PHM.70122
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 886-7466
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 29
; OTHER INFORMATION: /note= "Xaa is amidated leucine"
US-08-775-476-1

Query Match 57.7%; Score 41; DB 1; Length 29;
Best Local Similarity 54.5%; Pred. No. 31;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 CCSDRRCRXRC 12
||| ||| :
Db 15 CCEDMVCRLWC 25

RESULT 12
US-09-018-799-1
; Sequence 1, Application US/09018799
; Patent No. 5807821
; GENERAL INFORMATION:
; APPLICANT: Lampe, Richard A.
; TITLE OF INVENTION: Analgesic Peptides from Venom of
; TITLE OF INVENTION: Grammostola Spatulata and Use Thereof
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZENECA Pharmaceuticals
; STREET: 1800 Concord Pike
; CITY: Wilmington
; STATE: DE
; COUNTRY: United States
; ZIP: 19850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/018,799
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/775,476

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hohenschutz, Liza D.
; REGISTRATION NUMBER: 33,712
; REFERENCE/DOCKET NUMBER: PHM.70122
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 886-7466
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 29
; OTHER INFORMATION: /note= "Xaa is amidated leucine"
; US-09-018-799-1

Query Match 57.7%; Score 41; DB 1; Length 29;
Best Local Similarity 54.5%; Pred. No. 31;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 CCSDRRRCXRC 12
||| |||:
Db 15 CCEDMVCRLWC 25

RESULT 13
5189019-4
; Patent No. 5189019
; APPLICANT: LAMPE, RICHARD A.
; JWU-SHENG: LAW, SIMON W.; MARK, GEORGE E.
; TITLE OF INVENTION: ANALGESIC PEPTIDES FROM VENOM OF
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZENECA PHARMACEUTICALS
; STREET: 1800 CONCORD PIKE
; CITY: WILMINGTON
; STATE: DE
; COUNTRY: UNITED STATES
; ZIP: 19850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; OPERATING SYSTEM: IBM PC COMPATIBLE
; SOFTWARE: PATENT IN RELEASE #1.0, VERSION #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/099,974
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/018,799
; FILING DATE:
; APPLICATION NUMBER: US/08/775,476
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hohenschutz, Liza D.
; REGISTRATION NUMBER: 33,712
; REFERENCE/DOCKET NUMBER: PHM.70122
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 886-7466
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide

; LOCATION: 29
; OTHER INFORMATION: /note= "Xaa is amidated leucine"
; US-09-099-974-1

Query Match 57.7%; Score 41; DB 2; Length 29;
Best Local Similarity 54.5%; Pred. No. 31;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 CCSDRRRCXRC 12
||| |||:
Db 15 CCEDMVCRLWC 25

RESULT 14
5189019-8
; Patent No. 5189019
; APPLICANT: PALLADINO, LINDA O.; SILBERKLANG, MELVIN; TUNG,
; JWU-SHENG: LAW, SIMON W.; MARK, GEORGE E.
; TITLE OF INVENTION: ANTISTASIN DERIVED ANTICOAGULANT PROTEIN
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/512,695
; FILING DATE: 23-APR-1990
; SEQ ID NO: 8
; LENGTH: 58
; 5189019-8

Query Match 57.7%; Score 41; DB 6; Length 58;
Best Local Similarity 60.0%; Pred. No. 53;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 CSDRRRCXRC 12
||| |||:
Db 28 CSEVRCRVHC 37

RESULT 15
5189019-4
; Patent No. 5189019
; APPLICANT: PALLADINO, LINDA O.; SILBERKLANG, MELVIN; TUNG,
; JWU-SHENG: LAW, SIMON W.; MARK, GEORGE E.
; TITLE OF INVENTION: ANTISTASIN DERIVED ANTICOAGULANT PROTEIN
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/512,695
; FILING DATE: 23-APR-1990
; SEQ ID NO: 4
; LENGTH: 75
; 5189019-4

Query Match 57.7%; Score 41; DB 6; Length 75;
Best Local Similarity 60.0%; Pred. No. 64;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 CSDRRRCXRC 12
||| |||:
Db 45 CSEVRCRVHC 54

Search completed: July 1, 2002, 12:25:19
Job time: 156 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 12:23:58 ; Search time 16.34 Seconds
(without alignments)
28.435 Million cell updates/sec

Title: US-09-493-795A-5
Perfect score: 71
Sequence: 1 ACCSDRRRCXRC 12

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

arched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query | Score | Match | Length | DB | ID | Description |
|------------|-------|-------|-------|--------|------|-------|--------------------|
| 1 | 54 | 76.1 | 12 | 1 | CR22 | CR22 | P50983 conus imper |
| 2 | 43 | 60.6 | 202 | 1 | CR22 | CR22 | P23252 hordeum vul |
| 3 | 43 | 60.6 | 572 | 1 | CR22 | CR22 | Q19345 caenorhabdi |
| 4 | 42 | 59.2 | 1746 | 1 | TENA | PIG | Q29116 sus scrofa |
| 5 | 41 | 57.7 | 16 | 1 | CR22 | CR22 | P56638 conus epic |
| 6 | 41 | 57.7 | 69 | 1 | CR22 | CR22 | Q92286 conus texti |
| 7 | 41 | 57.7 | 204 | 1 | Y115 | NPVAC | P11219 oryza sativ |
| 8 | 41 | 57.7 | 227 | 1 | AGI | ORYSA | P34704 caenorhabdi |
| 9 | 40.5 | 57.0 | 175 | 1 | HER1 | CAEEL | P11219 oryza sativ |
| 10 | 40 | 56.3 | 396 | 1 | Y858 | TREPA | O83830 treponema p |
| 11 | 40 | 56.3 | 415 | 1 | TNR3 | MOUSE | P50284 mus musculu |
| 12 | 40 | 56.3 | 595 | 1 | P2X7 | MOUSE | Q92100 mus musculu |
| 13 | 39 | 54.9 | 46 | 1 | AX2 | BETVU | P82010 beta vulgar |
| 14 | 39 | 54.9 | 47 | 1 | STAL | SORBI | P21923 sorghum bic |
| 15 | 39 | 54.9 | 47 | 1 | TH22 | MAIZE | P81009 zea mays (m |
| 16 | 39 | 54.9 | 64 | 1 | MT | STPNE | P55953 sterechinus |
| 17 | 39 | 54.9 | 65 | 1 | MTB | STRPU | Q27287 strongyloce |
| 18 | 39 | 54.9 | 80 | 1 | TX3A | PHONI | P81793 phontutria |
| 19 | 39 | 54.9 | 82 | 1 | TX32 | PHONI | O76201 phontutria |
| 20 | 39 | 54.9 | 88 | 1 | CR14 | HORVU | P26154 hordeum vul |
| 21 | 39 | 54.9 | 109 | 1 | RL23 | AOUVY | Q92148 aquifex pyr |
| 22 | 39 | 54.9 | 202 | 1 | T451 | HUMAN | P30408 homo sapien |
| 23 | 39 | 54.9 | 257 | 1 | VG03 | HSVEB | P28988 equine herp |
| 24 | 39 | 54.9 | 258 | 1 | VG03 | HSVEB | P29124 equine herp |
| 25 | 39 | 54.9 | 497 | 1 | SC14 | YARLI | P45816 yarrowia li |
| 26 | 39 | 54.9 | 575 | 1 | FRPB | MYCTU | Q10547 mycobacteri |
| 27 | 39 | 54.9 | 769 | 1 | AD11 | HUMAN | O75078 homo sapien |
| 28 | 39 | 54.9 | 773 | 1 | AD11 | MOUSE | Q91144 mus musculu |
| 29 | 39 | 54.9 | 1013 | 1 | PRML | DROME | P82295 drosophila |
| 30 | 39 | 54.9 | 2034 | 1 | FER1 | CAEEL | Q17388 caenorhabdi |
| 31 | 39 | 54.9 | 2652 | 1 | RRPB | IBVB | P26314 avian infec |
| 32 | 38 | 53.5 | 50 | 1 | HSP1 | PONPY | P35310 pongo pygma |
| 33 | 38 | 53.5 | 55 | 1 | ANTA | HIRME | P80302 hirudo medi |

| | | | | | | | |
|----|----|------|------|---|------|-------|--------------------|
| 34 | 38 | 53.5 | 97 | 1 | KRB2 | CAPHI | P02447 capra hircu |
| 35 | 38 | 53.5 | 97 | 1 | KRB2 | SHEEP | P02446 ovis aries |
| 36 | 38 | 53.5 | 101 | 1 | THST | THETS | Q9nbw4 theromyzon |
| 37 | 38 | 53.5 | 136 | 1 | ANTA | HAEOF | P15358 haementeria |
| 38 | 38 | 53.5 | 231 | 1 | FLGH | PSEPU | Q52081 pseudomonas |
| 39 | 38 | 53.5 | 288 | 1 | YLG9 | CAEEL | P34412 caenorhabdi |
| 40 | 38 | 53.5 | 372 | 1 | AGI | URTDI | P11218 urtica dioi |
| 41 | 38 | 53.5 | 555 | 1 | FRPB | MYCLE | O33064 mycobacteri |
| 42 | 38 | 53.5 | 1224 | 1 | COPA | BOVIN | Q27954 bos taurus |
| 43 | 38 | 53.5 | 1224 | 1 | COPA | HUMAN | P53621 homo sapien |
| 44 | 38 | 53.5 | 1232 | 1 | KF4A | HUMAN | O95239 homo sapien |
| 45 | 38 | 53.5 | 2201 | 1 | TENA | HUMAN | P24821 homo sapien |

ALIGNMENTS

| RESULT | ID | Query | Match | Length | DB | ID | Description |
|--------|----|---|-------|--------|----------|-----|-------------|
| 1 | AC | 01-OCT-1996 (Rel. 34, Created) | 12 | AA | STANDARD | PRT | 12 AA |
| 2 | DT | 01-OCT-1996 (Rel. 34, Last sequence update) | 12 | AA | STANDARD | PRT | 12 AA |
| 3 | DT | 16-OCT-2001 (Rel. 40, Last annotation update) | 12 | AA | STANDARD | PRT | 12 AA |
| 4 | DE | Alpha-conotoxin Imi. | 12 | AA | STANDARD | PRT | 12 AA |
| 5 | OS | Conus Imperialis (Imperial cone). | 12 | AA | STANDARD | PRT | 12 AA |
| 6 | OC | Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda; | 12 | AA | STANDARD | PRT | 12 AA |
| 7 | OC | Neogastropoda; Conoidea; Conidae; Conus. | 12 | AA | STANDARD | PRT | 12 AA |
| 8 | OX | NCBI_TaxID=35631; | 12 | AA | STANDARD | PRT | 12 AA |
| 9 | RN | [1] | 12 | AA | STANDARD | PRT | 12 AA |
| 10 | RP | SEQUENCE, AND SYNTHESIS. | 12 | AA | STANDARD | PRT | 12 AA |
| 11 | RC | TISSUE-Venom: | 12 | AA | STANDARD | PRT | 12 AA |
| 12 | RX | MEDLINE=94266889; PubMed=82066995; | 12 | AA | STANDARD | PRT | 12 AA |
| 13 | RA | McIntosh J.M., Yoshikami D., Mane E., Nielsen D.B., Rivier J.E., | 12 | AA | STANDARD | PRT | 12 AA |
| 14 | RA | Gray W.R., Olivera B.M.; | 12 | AA | STANDARD | PRT | 12 AA |
| 15 | RT | "A nicotinic acetylcholine receptor ligand of unique specificity, | 12 | AA | STANDARD | PRT | 12 AA |
| 16 | RT | alpha-conotoxin Imi."; | 12 | AA | STANDARD | PRT | 12 AA |
| 17 | RL | J. Biol. Chem. 269:16733-16739(1994). | 12 | AA | STANDARD | PRT | 12 AA |
| 18 | RN | [2] | 12 | AA | STANDARD | PRT | 12 AA |
| 19 | RP | CHARACTERIZATION. | 12 | AA | STANDARD | PRT | 12 AA |
| 20 | RX | MEDLINE=95379776; PubMed=7651351; | 12 | AA | STANDARD | PRT | 12 AA |
| 21 | RA | Johnson D.S., Martinez J., Elgoyhen A.B., Heinemann S.F., | 12 | AA | STANDARD | PRT | 12 AA |
| 22 | RA | McIntosh J.M.; | 12 | AA | STANDARD | PRT | 12 AA |
| 23 | RT | "Alpha-conotoxin Imi exhibits subtype-specific nicotinic | 12 | AA | STANDARD | PRT | 12 AA |
| 24 | RT | acetylcholine receptor blockade: preferential inhibition of homomeric | 12 | AA | STANDARD | PRT | 12 AA |
| 25 | RT | alpha 7 and alpha 9 receptors."; | 12 | AA | STANDARD | PRT | 12 AA |
| 26 | RL | Mol. Pharmacol. 48:194-199(1995). | 12 | AA | STANDARD | PRT | 12 AA |
| 27 | RN | [3] | 12 | AA | STANDARD | PRT | 12 AA |
| 28 | RP | STRUCTURE BY NMR. | 12 | AA | STANDARD | PRT | 12 AA |
| 29 | RX | MEDLINE=99212205; PubMed=10194298; | 12 | AA | STANDARD | PRT | 12 AA |
| 30 | RA | Rogers J.P., Loguinbuhl P., Shen G.S., McCabe R.T., Stevens R.C., | 12 | AA | STANDARD | PRT | 12 AA |
| 31 | RA | Wemmer D.E.; | 12 | AA | STANDARD | PRT | 12 AA |
| 32 | RT | "NMR solution structure of alpha-conotoxin Imi and comparison to | 12 | AA | STANDARD | PRT | 12 AA |
| 33 | RT | other conotoxins specific for neuronal nicotinic acetylcholine | 12 | AA | STANDARD | PRT | 12 AA |
| 34 | RT | receptors."; | 12 | AA | STANDARD | PRT | 12 AA |
| 35 | RL | Biochemistry 38:3874-3882(1999). | 12 | AA | STANDARD | PRT | 12 AA |
| 36 | RN | [4] | 12 | AA | STANDARD | PRT | 12 AA |
| 37 | RP | STRUCTURE BY NMR. | 12 | AA | STANDARD | PRT | 12 AA |
| 38 | RX | MEDLINE=99158061; PubMed=10050774; | 12 | AA | STANDARD | PRT | 12 AA |
| 39 | RA | Methlenikov I.V., Sherkarev Z.O., Zhmak M.N., Ivanov V.T., | 12 | AA | STANDARD | PRT | 12 AA |
| 40 | RA | Methfessel C., Tsetlin V.I., Arseniev A.S.; | 12 | AA | STANDARD | PRT | 12 AA |
| 41 | RT | "NMR spatial structure of alpha-conotoxin Imi reveals a common | 12 | AA | STANDARD | PRT | 12 AA |
| 42 | RT | scaffold in snail and snake toxins recognizing neuronal nicotinic | 12 | AA | STANDARD | PRT | 12 AA |
| 43 | RT | acetylcholine receptors."; | 12 | AA | STANDARD | PRT | 12 AA |
| 44 | RL | FEBS Lett. 444:275-280(1999). | 12 | AA | STANDARD | PRT | 12 AA |
| 45 | RN | [5] | 12 | AA | STANDARD | PRT | 12 AA |
| 46 | RP | STRUCTURE BY NMR. | 12 | AA | STANDARD | PRT | 12 AA |
| 47 | RX | MEDLINE=99324017; PubMed=10395477; | 12 | AA | STANDARD | PRT | 12 AA |
| 48 | RA | Gehrmann J., Daly N.L., Alewood P.F., Craik D.J.; | 12 | AA | STANDARD | PRT | 12 AA |
| 49 | RT | "Solution structure of alpha-conotoxin Imi by 1H nuclear magnetic | 12 | AA | STANDARD | PRT | 12 AA |
| 50 | RT | resonance."; | 12 | AA | STANDARD | PRT | 12 AA |
| 51 | RL | J. Med. Chem. 42:2364-2372(1999). | 12 | AA | STANDARD | PRT | 12 AA |
| 52 | CC | - - FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY | 12 | AA | STANDARD | PRT | 12 AA |

CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM. IT IS HIGHLY ACTIVE AGAINST THE NEUROMUSCULAR
CC RECEPTOR IN FROG BUT NOT IN MICE. IN CONTRAST, IT INDUCES SEIZURES
CC WHEN INJECTED CENTRALLY IN MICE AND RATS. IT TARGET NEURONAL
CC NACHRS IN MAMMALS. BLOCKS HOMOMERIC ALPHA-7 NICOTINIC RECEPTORS
CC WITH THE HIGHEST APPARENT AFFINITY AND HOMOMERIC ALPHA-9 RECEPTORS
CC WITH 8-FOLD LOWER AFFINITY. HAS NO EFFECT ON RECEPTORS COMPOSED OF
CC ALPHA-3/BETA-2, ALPHA-3/BETA-2, ALPHA-4/BETA-2, ALPHA-2/BETA-4,
CC ALPHA-3/BETA-4, OR ALPHA-4/BETA-4 SUBUNIT COMBINATIONS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE ALPHA-TYPE CONOTOXIN FAMILY.
DR PDB; 1IM1; 15-JUN-99.
DR PDB; 1IMI; 23-APR-99.
DR PDB; 1CNI; 27-MAY-99.
KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
KW Venom; 3D-structure. 8
FT DISULFID 2 8
FT MOD_RES 3 12
FT 12 12 AMIDATION.
SQ SEQUENCE 12 AA; 1357 MW; 9C29CEA545A4176A CRC64;

Query Match 76.1%; Score 54; DB 1; Length 12;
Best Local Similarity 72.7%; Pred. No. 0.09;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CSDRRRCRXC 12
|||||:|
DB 2 CCSDPRCAWRC 12

RESULT 2
ID CR2_HORVU STANDARD; PRT; 202 AA.
AC P23252;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DE 01-NOV-1991 (Rel. 20, Last annotation update)
DE Cold-regulated protein 2 (Fragment).
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, GEORGIE;
RA Cattivelli L., Bartels D.;
RT "Molecular cloning and characterization of cold regulated genes in
barley.";
PL Plant Physiol. 93:1504-1510(1990).
CC -1- MISCELLANEOUS: CONTAINS SEVERAL ARGinine RESIDUES IN CLOSE
PROXIMITY WHICH MAY BE INVOLVED IN PROTEIN RNA INTERACTIONS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M60733; AAC98703.1; ALT_INIT.
DR EMBL; B45512; B45512.
DR PIR; B45512; B45512.
FT DOMAIN 1 125 ARG-RICH.
FT NON_TER 1
SQ SEQUENCE 202 AA; 21753 MW; 49A80A2D2D2B5C3E CRC64;

Query Match 60.6%; Score 43; DB 1; Length 202;
Best Local Similarity 77.8%; Pred. No. 19;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 CSDRRRCRXC 11
|||||:|
DB 112 CSDTRCGR 120

RESULT 3
ID NH25_CAEEL STANDARD; PRT; 572 AA.
AC Q19345; Q9NJ95; Q9NJ96; Q9NJ97;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Nuclear hormone receptor family member nhr-25.
GN NHR-25 OR F11C1.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Palmer S.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND CHARACTERIZATION.
RC STRAIN=BRISTOL N2;
RX MEDLINE=20237592; PubMed=10772806;
RA Gissendanner C.R., Sluder A.E.;
RT "nhr-25, the Caenorhabditis elegans ortholog of ftz-f1, is required
for epidermal and somatic gonad development.";
RL Dev. Biol. 221:259-272(2000).
CC -1- FUNCTION: Orphan nuclear receptor. Seems to be required during
development. Disruption of nhr-25 function leads to embryonic
arrest due to failure of the epidermally mediated process of
embryo elongation.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; alpha (shown here) and beta; are
produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: Expressed in the epidermis, the developing
somatic gonad, and a subset of other epithelial cells.
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z54270; CAA91028.1; -.
DR EMBL; AF179214; AAF67038.1; -.
DR EMBL; AF179215; AAF67039.1; -.
DR EMBL; AF179216; AAF67040.1; -.
DR HSSP; P19793; 2NLL.
DR WormPep; F11C1.6; CE03191.
DR InterPro; IPR000536; Hormone_rec_lig.
DR InterPro; IPR001628; zf-C4.
DR Pfam; PF00104; hormone_rec; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00047; STROIDFINGER.
DR SMART; SM00430; HOL1; 1.
DR SMART; SM00399; Znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger; Alternative splicing; Developmental protein.
FT DNA_BIND 18 83 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 18 38 C4-TYPE.
FT ZN_FING 54 78 C4-TYPE.
FT VARSPLIC 1 80 MISSING (IN ISOFORM BETA).
FT CONFLICT 100 100 F -> L (IN REF. 2; AAF67039).
FT CONFLICT 540 540 R -> Q (IN REF. 2; AAF67039).
SQ SEQUENCE 572 AA; 64772 MW; D3092AB533092D7A CRC64;

FT DISULFID 394 403 BY SIMILARITY.
 FT DISULFID 408 418 BY SIMILARITY.
 FT DISULFID 412 423 BY SIMILARITY.
 FT DISULFID 425 434 BY SIMILARITY.
 FT DISULFID 439 449 BY SIMILARITY.
 FT DISULFID 443 454 BY SIMILARITY.
 FT DISULFID 456 465 BY SIMILARITY.
 FT DISULFID 470 480 BY SIMILARITY.
 FT DISULFID 474 485 BY SIMILARITY.
 FT DISULFID 487 496 BY SIMILARITY.
 FT DISULFID 501 511 BY SIMILARITY.
 FT DISULFID 505 516 BY SIMILARITY.
 FT DISULFID 518 527 BY SIMILARITY.
 FT DISULFID 532 542 BY SIMILARITY.
 FT DISULFID 536 547 BY SIMILARITY.
 FT DISULFID 549 558 BY SIMILARITY.
 FT DISULFID 563 573 BY SIMILARITY.
 FT DISULFID 567 578 BY SIMILARITY.
 FT DISULFID 580 589 BY SIMILARITY.
 FT DISULFID 594 604 BY SIMILARITY.
 FT DISULFID 598 609 BY SIMILARITY.
 FT DISULFID 611 620 BY SIMILARITY.
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 788 788 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1034 1034 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1079 1079 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1121 1121 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1354 1354 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 1072 1162 MISSING (IN MAJOR ISOFORM).
 FT VARSPLIC 1072 1253 MISSING (IN MINOR-1 ISOFORM).
 FT CONFLICT 1007 1007 T -> M (IN REF. 2).
 SQ SEQUENCE 1746 AA; 191399 MW; 56549B1CFE5E5C88 CRC64;

Query Match 59.2%; Score 42; DB 1; Length 1746;
 Best Local Similarity 50.0%; Pred. No. 1.2e+02;
 Matches 9; Conservative 2; Mismatches 1; Indels 6; Gaps 1;

QY 1 ACSDRRRC-----RXRC 12
 DB 370 ADCSERRCPDCHNRGRC 387

RESULT 5
 CXAL_CONEP STANDARD; PRT; 16 AA.
 AC P56638;
 15-DEC-1998 (Rel. 37, Created)
 15-DEC-1998 (Rel. 37, Last sequence update)
 16-OCT-2001 (Rel. 40, Last annotation update)

DE Alpha-conotoxin Epi.
 OS Conus episcopatus (Bishop's cone).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.
 OX NCBI_TaxID=88764;
 RN [1]

RP X-RAY CRYSTALLOGRAPHY (1.1 ANGSTROMS).
 RX MEDLINE=98376423; PubMed=9708977;
 RA Hu S.H., Loughnan M., Miller R., Weeks C.M., Blessing R.H.,
 RA Alewood P.F., Lewis R.J., Martin J.L.;
 RT "The 1.1-A resolution crystal structure of [Tyr15]EpI, a novel
 alpha-conotoxin from Conus episcopatus, solved by direct methods.";
 RL Biochemistry 37:11425-11433(1998).
 CC !- FUNCTION: ALPHACONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
 CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
 CC INHIBIT THEM. THIS PEPTIDE BLOCKS MAMMALIAN NICOTINIC
 CC ACETYLCHOLINE RECEPTORS COMPOSED OF ALPHA-3/BETA-2 AND ALPHA-
 CC 3/BETA-4 SUBUNITS.
 CC !- SUBCELLULAR LOCATION: Secreted.
 CC !- SIMILARITY: BELONGS TO THE ALPHA-TYPE CONOTOXIN FAMILY.

DR PDB: 1AOM; 13-JAN-99.
 KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
 KW Venom; Sulfation; 3D-structure.
 FT DISULFID 2 8
 FT DISULFID 3 16
 FT MOD_RES 15 15 SULFATION.
 FT MOD_RES 16 16 AMIDATION.
 SQ SEQUENCE 16 AA; 1792 MW; C63385F376C99B4C CRC64;

Query Match 57.7%; Score 41; DB 1; Length 16;
 Best Local Similarity 66.7%; Pred. No. 5.1;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCSDRRCRX 10
 DB 2 CCSDPRCNM 10

RESULT 6
 CXAL_CONTE STANDARD; PRT; 69 AA.
 ID Q9X2K6;
 AC Q9X2K6;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alpha-type conotoxin Tx1 precursor.
 OS Conus textile (Cloth-of-gold cone).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.
 OX NCBI_TaxID=6494;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lu B.S., Yu F., Huang P.T., Huang C.F.;
 RT "RefSeq sequences of conotoxins from Conus textile.";
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

CC !- FUNCTION: ALPHACONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
 CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
 CC INHIBIT THEM (BY SIMILARITY).
 CC !- SUBCELLULAR LOCATION: Secreted.
 CC !- SIMILARITY: BELONGS TO THE ALPHA-TYPE CONOTOXIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF146352; AAC31912.1;
 KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Venom;
 KW Signal.
 FT SIGNAL 1 21 POTENTIAL.
 FT PROPEP 22 48 POTENTIAL.
 FT PEPTIDE 49 69 ALPHA-TYPE CONOTOXIN TX1.
 FT DISULFID 51 57 BY SIMILARITY.
 FT DISULFID 52 65 BY SIMILARITY.
 SQ SEQUENCE 69 AA; 7442 MW; E36ECE90BF1B56B0 CRC64;

Query Match 57.7%; Score 41; DB 1; Length 69;
 Best Local Similarity 66.7%; Pred. No. 15;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCSDRRCRX 10
 DB 51 CCSDPRCNS 59

RESULT 7
 ID Y115_NPVAC STANDARD; PRT; 204 AA.
 ID Y115_NPVAC

AC P41668;
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE Hypothetical 23.0 kDa protein in HE65-PK2 intergenic region.
 OS Autographa californica nuclear polyhedrosis virus (AcMNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=46015;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C6;
 RX MEDLINE=94303173; PubMed=8030224;
 RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
 RT "The complete DNA sequence of Autographa californica nuclear
 polyhedrosis virus";
 RT Virology 202:586-605(1994).
 BL -!- SIMILARITY: TO CORRESPONDING ORF IN OPMNPV.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).

 DR EMBL; L22858; AAA66745.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 204 AA; 23019 MW; 2EA2A52874C9A4AF CRC64;

 QY 2 CCSDRRCXRC 12
 | : : : : :
 DB 64 CVDYDQCRDNC 74
 | : : : : :

 RESULT 8
 AGI_ORYSA STANDARD; PRT; 227 AA.
 NC P11219;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DE Lectin precursor (Agglutinin).
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=CV, LEWONT;
 RX MEDLINE=92404719; PubMed=2535550;
 RA Wilkins T.A., Raikhel N.V.;
 RT "Expression of rice lectin is governed by two temporally and
 RT spatially regulated mRNAs in developing embryos";
 RL Plant Cell 1:541-549(1989).
 RN [2]
 RP PRELIMINARY SEQUENCE OF 123-157.
 RA Chapot M.-P., Peumans W.J., Strosberg A.D.;
 RT "Extensive homologies between lectins from non-leguminous plants.";
 RL FEBS Lett. 195:231-234(1986).
 CC -!- FUNCTION: N-ACETYL-D-GLUCOSAMINE BINDING LECTIN.
 CC -!- SIMILARITY: CONTAINS 4 CHITIN-BINDING DOMAINS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).

 DR EMBL; M24504; AAA20873.1; -;
 DR PIR; JQ1102; LNRZ.
 DR PIR; A23616; A23616.
 DR HSSP; P10968; 2CWG.
 DR InterPro; IPR001002; Chitin_bind.
 DR Pfam; PF00187; chitin_binding; 4.
 DR PRINTS; PR00451; CHITINBINDING.
 DR ProDom; PD000609; Chitin_bind; 3.
 DR SMART; SM00270; ChtBD1; 4.
 DR PROSITE; PS00026; CHITIN_BINDING; 4.
 KW Lectin; Repeat; Chitin-binding; Glycoprotein; Signal.
 FT SIGNAL 1 28
 FT CHAIN 29 201 LECTIN.
 FT PROPEP 202 227
 FT CHAIN 29 122 10 KDA PEPTIDE.
 FT CHAIN 123 201 8 KDA PEPTIDE.
 FT DOMAIN 29 72 CHITIN-BINDING 1.
 FT DOMAIN 73 115 CHITIN-BINDING 2.
 FT DOMAIN 116 158 CHITIN-BINDING 3.
 FT DOMAIN 159 201 CHITIN-BINDING 4.
 FT MOD_RES 29 29 PYRROLIDONE CARBOXYLIC ACID.
 FT DISULFID 31 46 BY SIMILARITY.
 FT DISULFID 40 52 BY SIMILARITY.
 FT DISULFID 45 59 BY SIMILARITY.
 FT DISULFID 63 68 BY SIMILARITY.
 FT DISULFID 74 89 BY SIMILARITY.
 FT DISULFID 83 95 BY SIMILARITY.
 FT DISULFID 88 102 BY SIMILARITY.
 FT DISULFID 106 111 BY SIMILARITY.
 FT DISULFID 117 132 BY SIMILARITY.
 FT DISULFID 126 138 BY SIMILARITY.
 FT DISULFID 131 145 BY SIMILARITY.
 FT DISULFID 149 154 BY SIMILARITY.
 FT DISULFID 160 175 BY SIMILARITY.
 FT DISULFID 169 181 BY SIMILARITY.
 FT DISULFID 174 188 BY SIMILARITY.
 FT DISULFID 192 197 BY SIMILARITY.
 FT CARBOHYD 211 211 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 227 AA; 22795 MW; 691EB39F6690BAF1 CRC64;

 Query Match 57.7%; Score 41; DB 1; Length 227;
 Best Local Similarity 75.0%; Pred. No. 37;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ACCSDRRC 8
 | : : : : :
 DB 67 ACCSQRC 74
 | : : : : :

 RESULT 9
 HERL_CAEEL STANDARD; PRT; 175 AA.
 ID HERL_CAEEL
 AC F34704;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE HER-1 protein precursor.
 GN HER-1 OR 2K287.8.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=93170661; PubMed=8436294;
 RA Perry M.D., Li W., Trent C., Robertson B., Fire A., Hageman J.M.,

DR EMBL; L38423; AAB00846.1; -;
 DR EMBL; U30798; AAA81334.1; -;
 DR HSSP; P25942; 1CDF.
 DR MGD; MGI:104875; Ltbr.
 DR InterPro; IPR001368; TNFR_c6.
 DR Pfam; PF00020; TNFR_c6; 3.
 DR ProDom; PD00071; TNFR_c6; 1.
 DR SMART; SM00208; TNFR; 3.
 DR PROSITE; PS00652; TNFR_NGFR_1; 2;
 DR PROSITE; PS00500; TNFR_NGFR_2; 3.
 KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
 FT SIGNAL 1 30 POTENTIAL.
 FT CHAIN 31 415 LYMPHOTOXIN-BETA RECEPTOR.
 FT DOMAIN 31 223 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 224 244 POTENTIAL.
 FT DOMAIN 245 415 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 42 81 TNFR-CYS 1.
 FT REPEAT 82 124 TNFR-CYS 2.
 FT REPEAT 125 170 TNFR-CYS 3.
 FT REPEAT 171 213 TNFR-CYS 4.
 FT DISULFID 43 58 BY SIMILARITY.
 FT DISULFID 59 72 BY SIMILARITY.
 FT DISULFID 62 80 BY SIMILARITY.
 FT DISULFID 83 98 BY SIMILARITY.
 FT DISULFID 101 116 BY SIMILARITY.
 FT DISULFID 104 124 BY SIMILARITY.
 FT DISULFID 126 132 BY SIMILARITY.
 FT DISULFID 139 150 BY SIMILARITY.
 FT DISULFID 142 169 BY SIMILARITY.
 FT DISULFID 172 187 BY SIMILARITY.
 FT CARBOHYD 40 40 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 179 179 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 415 AA; 44956 MW; 29B326A566AEF661 CRC64;

Query Match 56.3%; Score 40; DB 1; Length 415;
 Best Local Similarity 54.5%; Pred. No. 77;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 CCSDRCRCX 12
 DB 116 CTSRRKRC 126
 | | | | | : | |
 RESULT 12
 AX2_MOUSE
 ID AX2_MOUSE STANDARD; PRT; 595 AA.
 AC Q92LM0;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE P2X purinoceptor 7 (ATP receptor) (P2X7) (Purinergic receptor)
 DE (P22 receptor).
 GN P2RX7 OR P2X7.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10990;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99065122; PubMed=9849870;
 RA Chessel I.P., Simon J., Hibell A.D., Michel A.D., Barnard E.A.,
 RA Humphrey P.P.;
 FT "Cloning and functional characterisation of the mouse P2X7 receptor.";
 RL FEBS Lett. 439:28-30(1998).
 CC -!- FUNCTION: RECEPTOR FOR ATP THAT ACTS AS A LIGAND GATED ION
 CC CHANNEL. RESPONSIBLE FOR ATP-DEPENDENT LYSIS OF MACROPHAGES
 CC THROUGH THE FORMATION OF MEMBRANE PORES PERMEABLE TO LARGE
 CC MOLECULES. COULD FUNCTION IN BOTH FAST SYNAPTIC TRANSMISSION AND
 CC THE ATP-MEDIATED LYSIS OF ANTIGEN-PRESENTING CELLS.
 CC -!- SUBUNIT: HOMO- OR HETEROPOLYMERS (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: BELONGS TO THE P2X RECEPTOR FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 DR EMBL; AJ009823; CAA08853.1; -;
 DR MGD; MGI:133957; P2rx7.
 DR InterPro; IPR001429; P2X_receptor.
 DR Pfam; PF00864; P2X_receptor; 1.
 DR PRINTS; PR01307; P2XRECEPTOR.
 DR PROSITE; PS01212; P2X RECEPTOR; 1.
 KW Ionic channel; Transmembrane; Ion transport; Receptor; Glycoprotein.
 FT DOMAIN 1 25 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 26 46 1 (POTENTIAL).
 FT DOMAIN 47 334 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 335 355 2 (POTENTIAL).
 FT DOMAIN 356 595 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 74 74 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 187 187 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 202 202 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 213 213 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 595 AA; 68422 MW; 77AC8CA524162313 CRC64;

Query Match 56.3%; Score 40; DB 1; Length 595;
 Best Local Similarity 66.7%; Pred. No. 1e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 CCSDRCRCX 10
 DB 129 CSDRRCKK 137
 | | | | | : | |
 RESULT 13
 AX2_BETVU
 ID AX2_BETVU STANDARD; PRT; 46 AA.
 AC P82D10; P81510;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Antifungal protein AX2.
 OS Beta vulgaris (Sugar beet).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllales; Caryophyllaceae; Chenopodiaceae; Beta.
 OX NCBI_TaxID=3555;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=CV. TURBO, AND CV. RHIZOR; TISSUE=Leaf;
 RX MEDLINE=9538713; PubMed=7655063;
 RA Krach K.M., Nielsen J.E., Nielsen K.K., Dreboldt S., Mikkelsen J.D.;
 FT "Characterization and localization of new antifungal cysteine-rich
 FT proteins from Beta vulgaris.";
 RL Mol. Plant Microbe Interact. 8:424-434(1995).
 CC -!- FUNCTION: STRONG INHIBITING ACTIVITY AGAINST C.BETICOILA AND OTHER
 CC FILAMENTOUS FUNGI. LITTLE OR NO EFFECT AGAINST BACTERIA.
 CC -!- TISSUE SPECIFICITY: LEAVES AND FLOWERS.
 CC -!- MASS SPECTROMETRY: MW=5193; MW ERR=3; METHOD=Electrospray.
 CC -!- SIMILARITY: BELONGS TO THE GAMMA-PUROTHIONIN FAMILY.
 CC HSSP; P20158; 1GFS.
 DR InterPro; IPR002118; Gamma-thionin.
 DR InterPro; IPR003614; Knott1.
 DR Pfam; PF00304; Gamma-thionin; 1.
 DR ProDom; PD002594; Gamma-thionin; 1.
 DR SMART; SM00505; Knott1; 1.
 KW Antibiotic; Fungicide.
 FT DISULFID 3 46 BY SIMILARITY.
 FT DISULFID 14 34 BY SIMILARITY.

FT DISULFID 20 40 BY SIMILARITY.
FT DISULFID 24 42 BY SIMILARITY.
SQ SEQUENCE 46 AA; 5185 MW; 9A536EE9E74B18A6 CRC64;

Query Match 54.9% Score 39; DB 1; Length 46;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCSDRRCRXRC 12
| | | | | : |
DB 13 ACFSDTNCQAC 24

RESULT 14

SIAL_SORBI STANDARD; PRT; 47 AA.
AC P21923;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Small protein inhibitor of insect alpha-amylases 1 (SI alpha-1).
OC Sorghum bicolor (Sorghum) (Sorghum vulgare).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Sorghum.
OX NCBI_TaxID=4558;
RN [1]
RP SEQUENCE.
RC STRAIN=CV. FRENCH RED; TISSUE=Seed;
RX MEDLINE=91138737; PubMed=1995329;
RA Bloch C. Jr., Richardson M.;
RT "A new family of small (5 kDa) protein inhibitors of insect alpha-amylases from seeds or sorghum (Sorghum bicolor (L) Moench) have sequence homologies with wheat gamma-purothionins.";
RL FEBS Lett. 279:101-104(1991).
RN [2]
RP REVISION TO 35.
RC TISSUE=Seed;
RX MEDLINE=95220349; PubMed=7705336;
RA Nitti G., Orru S., Bloch C. Jr., Morhy L., Marino G., Pucci P.;
RT "Amino acid sequence and disulphide-bridge pattern of three gamma-thionins from Sorghum bicolor.";
RL Eur. J. Biochem. 228:250-256(1995).
RN [3]

RP STRUCTURE BY NMR.
RX MEDLINE=95220349; PubMed=7705336;
RA Bloch C. Jr., Patel S.V., Zvelebil M.J.J.M., Carr M.O., Sadler P.J., Thornton J.M.;
RL Unpublished results, cited by:
Nitti G., Orru S., Bloch C. Jr., Morhy L., Marino G., Pucci P.;
Eur. J. Biochem. 228:250-256(1995).
CC -1- FUNCTION: INHIBITS STRONGLY THE ALPHA-AMYLASES FROM THE GUTS OF LOCUST AND COCKROACH, WEAKLY THE ENZYME OF A.ORYZAE AND HUMAN SALIVA, AND FAILS TO INHIBIT THE ALPHA-AMYLASES FROM PORCINE PANCREAS, BARLEY AND BACILLUS SPP.

CC -1- SIMILARITY: BELONGS TO THE GAMMA-PUROTHIONIN FAMILY.

DR PIR; S13931; S13931.
DR InterPro; IPR002118; Gamma-thionin.
DR InterPro; IPR003614; Knottl.
DR Pfam; PF00304; Gamma-thionin; 1.
DR ProDom; PD002594; Gamma-thionin; 1.
DR SMART; SM00505; Knottl; 1.
DR PROSITE; PS00940; GAMMA_THIONIN; 1.
KW Alpha-amylase inhibitor.
FT DISULFID 3 47
FT DISULFID 14 36
FT DISULFID 20 41
FT DISULFID 24 43
SQ SEQUENCE 47 AA; 5382 MW; 4F0810A702232A0E CRC64;

Query Match

54.9% Score 39; DB 1; Length 47;

Best Local Similarity 54.5%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 CCSDRRCRXRC 12
| | | | | : |
DB 14 CISDRLCNEC 24

RESULT 15

THZ2_MAIZE STANDARD; PRT; 47 AA.
AC P81009;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gamma-zeathionin 2.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE.
RC TISSUE=Seed;
RA Castro M.S., Fontes W., Morhy L., Bloch C. Jr.;
RT "Complete amino acid sequences of two gamma-thionins from maize (Zea mays L.) seeds.";
RL Protein Pept. Lett. 3:267-274(1996).
CC -1- FUNCTION: THIONINS ARE SMALL PLANT PROTEINS WHICH ARE TOXIC TO ANIMAL CELLS.
CC -1- SIMILARITY: BELONGS TO THE GAMMA-PUROTHIONIN FAMILY.
DR MaizeDB; 139775; -;
DR InterPro; IPR002118; Gamma-thionin.
DR InterPro; IPR003614; Knottl.
DR Pfam; PF00304; Gamma-thionin; 1.
DR ProDom; PD002594; Gamma-thionin; 1.
DR SMART; SM00505; Knottl; 1.
DR PROSITE; PS00940; GAMMA_THIONIN; 1.
KW Thionin; Plant toxin.
FT DISULFID 3 47
FT DISULFID 14 36
FT DISULFID 20 41
FT DISULFID 24 43
SQ SEQUENCE 47 AA; 5368 MW; 4F0810A6D9532A0E CRC64;

Query Match

54.9% Score 39; DB 1; Length 47;
Best Local Similarity 54.5%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 CCSDRRCRXRC 12
| | | | | : |
DB 14 CISDRLCNEC 24

Search completed: July 1, 2002, 12:27:03
Job time: 185 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 12:23:08 ; Search time 25.3 Seconds
(without alignments)
45.576 Million cell updates/sec

Title: US-09-493-795A-5
Perfect score: 71
Sequence: 1 ACCSDRRRCXRC 12

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71.*

- 1: pir1.*
- 2: pir2.*
- 3: pir3.*
- 4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------|---------------------|
| 1 | 54 | 76.1 | 12 | A53709 | alpha-conotoxin ImI |
| 2 | 47 | 66.2 | 1057 | T30638 | hypothetical prote |
| 3 | 46 | 64.8 | 174 | B95940 | hypothetical prote |
| 4 | 46 | 64.8 | 521 | T46250 | hypothetical prote |
| 5 | 46 | 64.8 | 3570 | T45025 | mucin MUC5B, trach |
| 6 | 45 | 63.4 | 496 | T42562 | host shutoff virio |
| 7 | 43 | 60.6 | 202 | B45512 | cold-regulated pro |
| 8 | 43 | 60.6 | 496 | T75257 | hypothetical prote |
| 9 | 43 | 60.6 | 572 | T20764 | hypothetical prote |
| 10 | 42 | 59.2 | 174 | T21945 | hypothetical prote |
| 11 | 42 | 59.2 | 1746 | S19694 | tenascin precursor |
| 12 | 41.5 | 58.5 | 183 | T34798 | hypothetical prote |
| 13 | 41 | 57.7 | 16 | A59042 | alpha-conotoxin Ep |
| 14 | 41 | 57.7 | 136 | JS0209 | antistasin precurs |
| 15 | 41 | 57.7 | 177 | S37650 | high-sulfur kerati |
| 16 | 41 | 57.7 | 183 | T23240 | hypothetical prote |
| 17 | 41 | 57.7 | 204 | T41853 | AcMNPV orf115 - Bo |
| 18 | 41 | 57.7 | 204 | D72864 | AcOrf-115 protein |
| 19 | 41 | 57.7 | 227 | LNR2 | lectin precursor - |
| 20 | 41 | 57.7 | 240 | E45714 | probable regulator |
| 21 | 41 | 57.7 | 241 | C46181 | pX-tax-orf II (alt |
| 22 | 41 | 57.7 | 1274 | T02636 | D1 protein homolog |
| 23 | 41 | 57.7 | 2825 | T14271 | Doc4 protein, stre |
| 24 | 40.5 | 57.0 | 175 | A46388 | Her-1 protein - Ca |
| 25 | 40 | 56.3 | 186 | A45910 | ultra-high-sulfur |
| 26 | 40 | 56.3 | 380 | T04508 | hypothetical prote |
| 27 | 40 | 56.3 | 408 | B71272 | hypothetical prote |
| 28 | 40 | 56.3 | 2019 | JQ1322 | tenascin precursor |
| 29 | 39.5 | 55.6 | 336 | G84562 | hypothetical prote |

| | | | | | | |
|----|------|------|------|---|--------|----------------------|
| 30 | 39.5 | 55.6 | 609 | 2 | B86407 | auxin-regulated pr |
| 31 | 39.5 | 55.6 | 1126 | 2 | A96032 | probable two-compo |
| 32 | 39.5 | 55.6 | 1131 | 2 | A97650 | hypothetical prote |
| 33 | 39.5 | 55.6 | 1131 | 2 | A42873 | two component sens |
| 34 | 39 | 54.9 | 34 | 2 | B44336 | neurotoxin Trx3-2 - |
| 35 | 39 | 54.9 | 47 | 2 | B58319 | gamma-zeathlonin 2 |
| 36 | 39 | 54.9 | 47 | 2 | S69145 | metallothionein SI-a |
| 37 | 39 | 54.9 | 65 | 2 | A38739 | metallothionein - |
| 38 | 39 | 54.9 | 72 | 2 | A42325 | orf 5' to pheC - p |
| 39 | 39 | 54.9 | 88 | 2 | S16161 | BLT14 protein - ba |
| 40 | 39 | 54.9 | 93 | 2 | S60079 | oct2 protein isofo |
| 41 | 39 | 54.9 | 169 | 2 | T26271 | hypothetical prote |
| 42 | 39 | 54.9 | 202 | 2 | A42926 | L6 surface protein |
| 43 | 39 | 54.9 | 212 | 2 | C87585 | hypothetical prote |
| 44 | 39 | 54.9 | 257 | 1 | W2BEA2 | gene 3 protein - e |
| 45 | 39 | 54.9 | 258 | 1 | W2BEKA | gene 3 protein - e |

ALIGNMENTS

RESULT 1
A53709
alpha-conotoxin ImI - cone shell (Conus imperialis)
N:Alternate names: alpha-CTX-ImI
C:Species: Conus imperialis (imperial cone)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A53709
R:McIntosh, J.M.; Yoshikami, D.; Mahe, E.; Nielsen, D.B.; Rivier, J.E.; Gray, W.R.; O
J. Biol. Chem. 269, 16733-16739, 1994
A:Title: A nicotinic acetylcholine receptor ligand of unique specificity, alpha-conot
A:Reference number: A53709; MUID:94266889
A:Accession: A53709
A:Molecule type: protein
A:Residues: 1-12 <MC>
A:Note: structure confirmed by chemical synthesis
C:Comment: This alpha-conotoxin, as an acetylcholine receptor inhibitor, is a postsyn
C:Superfamily: alpha-conotoxin
C:Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; postsynaptic neu
F:2-8,3-12/Disulfide bonds: #status experimental
F:12/Modified site: amidated carboxyl end (Cys) #status experimental

Query Match 76.1%; Score 54; DB 1; Length 12;
Best Local Similarity 72.7%; Pred No. 0.32;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCSDRRRCXRC 12
 ||||| :||
DB 2 CCSDPRCAWRC 12

RESULT 2
T30638
hypothetical protein 36R - Mollusca contagiosum virus 1
N:Alternate names: MC036R
C:Species: Mollusca contagiosum virus 1
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 11-May-2000
C:Accession: T30638
R:Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Daral, G.; Moss, B.
Science 273, 813-816, 1996
A:Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host
A:Reference number: 220876; MUID:96325459
A:Accession: T30638
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1057 <SEN>
A:Cross-references: EMBL:U60315; PIDN:AAC55164.1
C:Genetics:
A:Note: MC036R

Query Match 66.2%; Score 47; DB 2; Length 1057;

Best Local Similarity 70.0%; Pred. NO. 40;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 CSDRRRCRXC 12
DB 895 CLDRCRRRC 904

RESULT 3
B95940
hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) megaplasmid pSymb
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: B95940
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,683-Kb pSymb megaplasmid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: B95940
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-174 <CUR>
A:Cross-references: GB:AL591985; PIDN:CAC49186.1; PID:gl5140671; GSPDB:GN00167
Experimental source: strain 1021, megaplasmid pSymb
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wellis, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: Smb21282
A:Genome: plasmid

Query Match 64.8%; Score 46; DB 2; Length 174;
Best Local Similarity 63.6%; Pred. No. 18;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCSDRRRCRXC 12
DB 44 CCLRRRCRLPC 54

RESULT 4
T46250
hypothetical protein DKF2p761A051.1 - human (fragment)
C:Species: Homo sapiens (nan)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T46250
R:Ottenwaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23031
A:Accession: T46250
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-521 <AAA>
A:Cross-references: EMBL:AL137496
A:Experimental source: adult amygdala; clone DKF2p761A051
C:Genetics:
A:Note: DKF2p761A051.1

Query Match 64.8%; Score 46; DB 2; Length 521;
Best Local Similarity 60.0%; Pred. No. 35;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACCSDRRRCRX 10
DB 440 SCCSDHRCRW 449

RESULT 5
T45025
mucin MUC5B, tracheobronchial [imported] - human (fragment)
C:Species: Homo sapiens (nan)
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C:Accession: T45025
R:Desseyn, J.L.; Guyonnet-Duport, V.; Porchet, N.; Aubert, J.P.; Laine, A.
J. Biol. Chem. 272, 3168-3178, 1997
A:Title: Human mucin gene MUC5B, the 10.7 kb large central exon encodes various alter
A:Reference number: Z22899; MUID:97166151
A:Accession: T45025
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-3570 <DES>
A:Cross-references: EMBL:Z72496; NID:gl834502; PIDN:CAA96577.1; PID:gl834503
A:Experimental source: placenta
C:Genetics:
A:Gene: MUC5B

Query Match 64.8%; Score 46; DB 2; Length 3570;
Best Local Similarity 70.0%; Pred. No. 11e-02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCSDRRRCRXC 11
DB 273 CCSDHRCRGR 282

RESULT 6
T42562
host shutoff virion protein - equine herpesvirus 4 (strain NS80567)
C:Species: equine herpesvirus 4
A:Variety: strain NS80567
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C:Accession: T42562
R:Telford, E.A.; Watson, M.S.; Perry, J.; Cullinane, A.A.; Davison, A.J.
J. Gen. Virol. 79, 1197-1203, 1998
A:Title: The DNA sequence of equine herpesvirus-4.
A:Reference number: Z22173; MUID:98264497
A:Accession: T42562
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-496 <TEL>
A:Cross-references: EMBL:AF030027; NID:g2605950; PIDN:AAC59533.1; PID:g2605961
A:Experimental source: strain NS80567
C:Genetics:
A:Note: 19
C:Superfamily: varicella-zoster virus gene 17 protein

Query Match 63.4%; Score 45; DB 2; Length 496;
Best Local Similarity 72.7%; Pred. No. 46;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCSDRRRCRXC 11
DB 343 ACLSDRGCRXR 353

RESULT 7
B45512
cold-regulated protein 2 - barley (fragment)
C:Species: Hordeum vulgare (barley)
C:Date: 03-Jun-1993 #sequence_revision 30-Sep-1993 #text_change 09-Sep-1997
C:Accession: B45512
R:Cattivelli, L.; Bartels, D.
Plant Physiol. 93, 1504-1510, 1990
A:Title: Molecular cloning and characterization of cold-regulated genes in barley.
A:Reference number: A45512
A:Accession: B45512

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-202 <CAT>
A:Cross-references: GB:M60733; NID:g167027; PID:g167028

Query Match 60.6%; Score 43; DB 2; Length 202;
Best Local Similarity 77.8%; Pred. No. 49;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 CSDRRRCRXR 11
||| |||:
DB 112 CSDTRCGR 120

RESULT 8

hypothetical protein - Deinococcus radiodurans (strain R1)
Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: F75257
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896

A:Accession: F75257

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-496 <WHI>

A:Cross-references: GB:AE002086; GB:AE000513; NID:g6460395; PIDN:AAF12116.1; PID:g646040

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR2572

A:Map position: 1

Query Match 60.6%; Score 43; DB 2; Length 496;
Best Local Similarity 63.6%; Pred. No. 85;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCSRRRCRXR 11
||| |||:
DB 485 ACCSRRCCCR 495

RESULT 9

hypothetical protein Fl1C1.6 - Caenorhabditis elegans
Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T20764
R:Palmer, S.
submitted to the EMBL Data Library, September 1995

A:Reference number: Z19321

A:Accession: T20764

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-572 <WIL>

A:Cross-references: EMBL:254270; PIDN:CAA91028.1; GSPDB:GN000028; CESP:Fl1C1.6

A:Experimental source: clone Fl1C1

C:Genetics:

A:Gene: CESP:Fl1C1.6

A:Map position: X

A:Introns: 39/3; 50/3; 87/1; 148/2; 190/1; 286/1; 377/3; 417/2; 499/2

C:Superfamily: steroid hormone receptor Ad4BP; erba transforming protein homology

Query Match 60.6%; Score 43; DB 2; Length 572;
Best Local Similarity 63.6%; Pred. No. 92;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCSDRRRCRXR 12
||| |||:
DB 60 CHVDTRCKRC 70

RESULT 10

T21945
hypothetical protein F38B2.2 - Caenorhabditis elegans
Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T21945
R:Kershaw, J.
submitted to the EMBL Data Library, July 1995

A:Reference number: Z19492

A:Accession: T21945

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-174 <WIL>

A:Cross-references: EMBL:250045; PIDN:CAA90362.1; GSPDB:GN000028; CESP:F38B2.2

A:Experimental source: clone F38B2

C:Genetics:

A:Gene: CESP:F38B2.2

A:Map position: X

A:Introns: 41/1; 81/3; 125/3

Query Match 59.2%; Score 42; DB 2; Length 174;
Best Local Similarity 54.5%; Pred. No. 61;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCSDRRRCRXR 12
||| |||:
DB 85 CCSEGCCLTRC 95

RESULT 11

tenascin precursor - pig
N:Alternate names: contactin; hexabrachion
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S19694
R:Nishi, T.; Weinstein, J.; Gillespie, W.M.; Paulson, J.C.
Eur. J. Biochem. 202, 643-648, 1991

A:Title: Complete primary structure of porcine tenascin. Detection of tenascin trans

A:Reference number: S19694; MUID:92104189

A:Accession: S19694

A:Molecule type: mRNA

A:Residues: 1-1746 <NIS>

A:Cross-references: EMBL:X61599; NID:g2124; PIDN:CAA43796.1; PID:g2125

C:Superfamily: tenascin; EGF homology; fibrinogen beta/gamma homology; fibronectin ty
C:Keywords: alternative splicing; calcium binding; cell adhesion; duplication; extrac
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-1746/Product: tenascin #status predicted <MAT>

F:346-372/Domain: EGF homology <EGF>

F:377-403/Domain: EGF homology <EGF>

F:622-703/Domain: fibronectin type III repeat homology <FN3A>

F:711-793/Domain: fibronectin type III repeat homology <FN3B>

F:802-884/Domain: fibronectin type III repeat homology <FN3C>

F:892-976/Domain: fibronectin type III repeat homology <FN3D>

F:984-1064/Domain: fibronectin type III repeat homology <FN3E>

F:1073-1155/Domain: fibronectin type III repeat homology <FN3F>

F:1164-1246/Domain: fibronectin type III repeat homology <FN3G>

F:1254-1335/Domain: fibronectin type III repeat homology <FN3H>

F:1343-1423/Domain: fibronectin type III repeat homology <FN3I>

F:1431-1511/Domain: fibronectin type III repeat homology <FN3J>

F:1526-1734/Domain: fibrinogen beta/gamma homology <FBG>

F:38.166.184.327.788.1034.1079.1121.1354/Binding site: carbohydrate (Asn) (covalent)

Query Match 59.2%; Score 42; DB 1; Length 1746;
Best Local Similarity 50.0%; Pred. No. 2.5e+02;
Matches 9; Conservative 2; Mismatches 1; Indels 6; Gaps 1;

QY 1 ACCSDRRRC-----RXRC 12
| ||||| |:
Db 370 ADCSERRCPSDCHNRGRC 387

RESULT 12

T34798
hypothetical protein SC2E1.30 - Streptomyces coelicolor

C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 04-Mar-2000
C:Accession: T34798

R:Murphy, L.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, June 1998

A:Reference number: Z21557

A:Accession: T34798

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-183 <MUR>

A:Cross-references: EMBL:AL023797; PIDN:CAA19405.1; GSPDB:GN00070; SCOEDB:SC2E1.30

A:Experimental source: strain A3(2)

C:Genetics:

Gene: SCOEDB:SC2E1.30

Superfamily: Streptomyces coelicolor hypothetical protein SC2E1.30

Query Match 58.5%; Score 41.5; DB 2; Length 183;
Best Local Similarity 42.1%; Pred. No. 73;
Matches 8; Conservative 1; Mismatches 3; Indels 7; Gaps 1;

QY 1 ACCS-----DRRCRXRC 12
||| ||| |:
Db 69 ACCPRWLTHKVSRRCSRC 87

RESULT 13

A59042

alpha-conotoxin Epi - cone shell (Conus episcopatus)

C:Species: Conus episcopatus (bishop's cone)

C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 13-Aug-1999

C:Accession: A59042

R:Loughnan, M.; Bond, T.; Atkins, A.; Cuevas, J.; Adams, D.J.; Broxton, N.M.; Livett, B.

J. Biol. Chem. 273, 15667-15674, 1998

A:Title: Alpha-conotoxin Epi, a novel sulfated peptide from Conus episcopatus that selected

A:Reference number: A59042; MUID:98288307

A:Accession: A59042

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-16 <LOU>

C:Superfamily: alpha-conotoxin

Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; postsynaptic neuroto

-16/Product: alpha-conotoxin Epi #status experimental <MAT>

-8,3-16/Disulfide bonds: #status experimental

F:15/Binding site: sulfate (tyr) (covalent) #status experimental

F:16/Modified site: amidated carboxyl end (cys) #status experimental

Query Match 57.7%; Score 41; DB 2; Length 16;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCSDRRRCRX 10

||||| |:
Db 2 CCSDPRCNM 10

RESULT 14

JS0209

antistatin precursor - Mexican leech

C:Species: Haementeria officinalis (Mexican leech)

C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 20-Aug-1999

C:Accession: JS0209

R:Han, J.H.; Law, S.W.; Keller, P.M.; Kniskern, P.J.; Silberklang, M.; Tung, J.S.; Gasid

Gene 75, 47-57, 1989
A:Title: Cloning and expression of cDNA encoding antistatin, a leech-derived protein
A:Reference number: JS0209; MUID:89252921

A:Accession: JS0209

A:Molecule type: mRNA

A:Residues: 1-136 <HAN>

A:Cross-references: GB:M24423; NID:gl59218; PIDN:AAA29193.1; PID:gl59219

A:Experimental source: clones lambda 0-12 and lambda 5C-4

A:Note: the clone lambda 5C-4 sequence is shown; the sequence of clone lambda 0-13 di

C:Comment: This protein exhibits anticoagulant and antimetastatic activity. It exerts

C:Superfamily: antistatin

C:Keywords: anticoagulant

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-136/Product: antistatin #status predicted <MAT>

F:18-71,72-127/Region: tandem repeats

Query Match 57.7%; Score 41; DB 2; Length 136;
Best Local Similarity 60.0%; Pred. No. 70;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 CSDRRRCRXRC 12

||| ||| |:
Db 45 CSEVRCRVHC 54

RESULT 15

S37650

high-sulfur keratin - human

C:Species: Homo sapiens (man)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Aug-1999

C:Accession: S37650

R:Zhunabava, B.D.; Gening, L.V.; Gazaryan, K.G.

Mol. Biol. 26, 550-555, 1992

A:Title: Cloning and structural characterization of human hair sulfur-rich keratin ge

A:Reference number: S37649

A:Accession: S37650

A:Molecule type: DNA

A:Residues: 1-177 <ZHU>

A:Cross-references: EMBL:X63337; NID:g311879; PIDN:CAA44937.1; PID:g311880

C:Superfamily: keratin high-sulfur matrix protein IIIA

Query Match 57.7%; Score 41; DB 2; Length 177;
Best Local Similarity 41.7%; Pred. No. 83;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCSDRRRCRXRC 12

||| ||| |:
Db 28 SCCESSCQPRC 39

Search completed: July 1, 2002, 12:25:52
Job time: 164 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 12:23:38 ; Search time 41.46 Seconds
(without alignments)
50.071 Million cell updates/sec

Title: US-09-493-795A-5
Perfect score: 71
Sequence: 1 ACCSDRRCXRC 12

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

arched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-----------|---------------------|
| 1 | 48 | 67.6 | 1516 | 4 Q12764 | Q12764 homo sapien |
| 2 | 48 | 67.6 | 4360 | 3 Q9UVN5 | Q9uvn5 alternaria |
| 3 | 47 | 66.2 | 228 | 12 Q11320 | Q11320 molluscum c |
| 4 | 47 | 66.2 | 1057 | 12 Q98204 | Q98204 molluscum c |
| 5 | 46 | 64.8 | 174 | 16 Q92VC2 | Q92vc2 rhizobium m |
| 6 | 46 | 64.8 | 399 | 12 Q69370 | Q69370 cercopithec |
| 7 | 46 | 64.8 | 488 | 4 Q9HAU7 | Q9ha77 homo sapien |
| 8 | 46 | 64.8 | 521 | 4 Q9NT71 | Q9nt71 homo sapien |
| 9 | 46 | 64.8 | 523 | 4 Q9HA72 | Q9hat2 homo sapien |
| 10 | 46 | 64.8 | 541 | 11 P70565 | P70565 mus musculus |
| 11 | 46 | 64.8 | 541 | 11 Q922L0 | Q922l0 mus musculus |
| 12 | 46 | 64.8 | 3570 | 4 Q99552 | Q99552 homo sapien |
| 13 | 45 | 63.4 | 496 | 12 Q39261 | Q39261 equine herp |
| 14 | 44 | 62.0 | 106 | 4 Q9BYU7 | Q9byu7 homo sapien |
| 15 | 44 | 62.0 | 193 | 4 Q9BYO5 | Q9byq5 homo sapien |
| 16 | 44 | 62.0 | 208 | 5 Q9W2F5 | Q9w2f5 drosophila |

| | | | | | |
|----|------|------|------|-----------|--------------------|
| 17 | 44 | 62.0 | 210 | 4 Q9BYR0 | Q9byr0 homo sapien |
| 18 | 44 | 62.0 | 520 | 5 Q25397 | Q25397 lynnaea sta |
| 19 | 43 | 60.6 | 496 | 16 Q9RRC0 | Q9rrc0 deinococcus |
| 20 | 43 | 60.6 | 1011 | 16 Q981K8 | Q981k8 rhizobium l |
| 21 | 42 | 59.2 | 128 | 11 Q925H4 | Q925h4 mus musculu |
| 22 | 42 | 59.2 | 154 | 4 Q9BYP9 | Q9byp9 homo sapien |
| 23 | 42 | 59.2 | 174 | 5 Q20138 | Q20138 caenorhabdi |
| 24 | 42 | 59.2 | 195 | 2 Q5JMW0 | Q5jmw0 bradyrhizob |
| 25 | 42 | 59.2 | 195 | 4 Q9BYQ6 | Q9byq6 homo sapien |
| 26 | 41.5 | 58.5 | 183 | 2 Q69901 | Q69901 streptomyce |
| 27 | 41 | 57.7 | 131 | 5 Q9W3A7 | Q9w3a7 drosophila |
| 28 | 41 | 57.7 | 177 | 4 Q07627 | Q07627 homo sapien |
| 29 | 41 | 57.7 | 183 | 5 Q21131 | Q21131 caenorhabdi |
| 30 | 41 | 57.7 | 200 | 10 Q9SBW5 | Q9sbw5 oryza sativ |
| 31 | 41 | 57.7 | 200 | 10 Q9XFF2 | Q9xff2 oryza sativ |
| 32 | 41 | 57.7 | 200 | 10 Q9XFF3 | Q9xff3 oryza sativ |
| 33 | 41 | 57.7 | 200 | 10 Q9XFF4 | Q9xff4 oryza sativ |
| 34 | 41 | 57.7 | 204 | 12 Q92472 | Q92472 bombyx mori |
| 35 | 41 | 57.7 | 240 | 15 Q9PXZ9 | Q9pxz9 human t-cel |
| 36 | 41 | 57.7 | 241 | 15 Q9PXV7 | Q9pxv7 human t-cel |
| 37 | 41 | 57.7 | 714 | 4 Q14526 | Q14526 homo sapien |
| 38 | 41 | 57.7 | 1274 | 5 Q60985 | Q60985 dictyosteli |
| 39 | 41 | 57.7 | 2771 | 11 Q9WTS7 | Q9wts7 mus musculu |
| 40 | 41 | 57.7 | 2825 | 11 Q70465 | Q70465 mus musculu |
| 41 | 41 | 57.7 | 3021 | 12 Q92933 | Q92933 hepatitis c |
| 42 | 41 | 57.7 | 3680 | 5 Q9VR08 | Q9vr08 drosophila |
| 43 | 41 | 57.7 | 6781 | 12 Q91AV2 | Q91av2 porcine epi |
| 44 | 40 | 56.3 | 59 | 5 P82107 | P82107 hirudo medl |
| 45 | 40 | 56.3 | 70 | 5 Q9BPJ6 | Q9bpj6 conus texti |

ALIGNMENTS

RESULT 1

Q12764 PRELIMINARY; PRT; 1516 AA.
AC Q12764;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE KIAA0191 PROTEIN (FRAGMENT).
GN KIAA0191
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Nagase T., Seki N., Ishikawa K., Tanaka A., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. V.
RT The coding sequences of 40 new genes (KIAA0161-KIAA0200) deduced
RT by analysis of cDNA clones from human cell line KG-1.";
RL DNA Res. 0:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96281124; PubMed=8724849;
RA Nagase T., Seki N., Ishikawa K., Tanaka A., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. V.
RT The coding sequences of 40 new genes (KIAA0161-KIAA0200) deduced by
RT analysis of cDNA clones from human cell line KG-1.";
RL DNA Res. 3:17-24(1996).
DR EMBL; D83776; BAA12105.1; -;
DR InterPro; IPR002934; NTP_transf.
DR InterPro; IPR001201; PAP_25A_core.
DR InterPro; IPR002058; PAP_assoc.
DR InterPro; IPR000822; Znf_C2H2.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF01909; NTP_transf_2; 1.
DR Pfam; PF00098; zf-CCHC; 3.
DR SMART; SM00343; Znf_C2HC; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW Zinc-finger.

FT NON_TER 1 1
SQ SEQUENCE 1516 AA; 171110 MW; 3FB7736E4FB8AE26 CRC64;

Query Match 67.6%; Score 48; DB 4; Length 1516;
Best Local Similarity 70.0%; Pred. No. 7.5;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 CSDRRRCRXC 12

I: |||||:|

Db 1497 CATRRRCRXC 1506

RESULT 2

Q9UVN5 PRELIMINARY; PRT; 4360 AA.

AC Q9UVN5; STRAIN-M-71;

DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE AM-TOXIN SYNTHETASE.

AMT.

Alternaria alternata.

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;

OC Pleosporales; Pleosporaceae; mitosporic Pleosporaceae; Alternaria.

OX NCBI_TaxID=5599;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-M-71;

RX MEDLINE=20331590; PubMed=10875335;

RA Johnson R.D., Johnson L., Itoh Y., Kodama M., Otani H., Kohmoto K.;

RT "Cloning and Characterization of a Cyclic Peptide Synthetase Gene from

RT Alternaria alternata Apple Pathotype Whose Product Is Involved in AM-

RT Toxin Synthesis and Pathogenicity";

RL Mol. Plant Microbe Interact. 13:742-753(2000).

DR EMBL; AF184074; AAF01762.1; -.

DR HSSP; P14687; 1AMD.

DR InterPro; IPR000873; AMP-bind.

DR InterPro; IPR001242; DUF4.

DR InterPro; IPR003015; HLH_MVC.

DR InterPro; IPR003880; Phosphopant_attach.

DR Pfam; PF00501; AMP-binding; 5.

DR Pfam; PF00668; Condensation; 6.

DR Pfam; PF00550; pp-binding; 4.

DR PROSITE; PS00075; ACP_DOMAIN; 4.

DR PROSITE; PS00455; AMP_BINDING; 3.

DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.

DR PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN_4.

DR Phosphopantetheine.

SEQUENCE 4360 AA; 479036 MW; A5172C653CEBA468 CRC64;

Query Match 67.6%; Score 48; DB 3; Length 4360;

Best Local Similarity 87.5%; Pred. No. 18;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCDRRC 8

|||:||||

Db 3044 ACCTDRRC 3051

RESULT 3

O11320

ID O11320

AC O11320

DT 01-JUL-1997 (TREMBlrel. 04, Created)

DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE HYPOTHETICAL 26.5 KDA PROTEIN (FRAGMENT).

GN B-N' 1.

OS Molluscum contagiosum virus subtype 1 (MCV1).

OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

OC Molluscipoxvirus.

OX NCBI_TaxID=10280;

RN [1]

RP SEQUENCE FROM N.A.

RA Moratilla M., Agramayor M., Nunez A., Funes J.M., Varas A.J.,

RA Lopez-Esteban J.L., Esteban M., Martin-Gallardo A.;

RT "A Random DNA Sequencing, Computer-Based Approach for the Generation

RT of a Gene Map of Molluscum Contagiosum Virus.";

RL Virus Genes 0:0-0(1997).

DR EMBL; U86902; AAB57950.1; -.

KW Hypothetical protein.

FT NON_TER 1

SQ SEQUENCE 228 AA; 26523 MW; E6321563A285F9B9 CRC64;

Query Match 66.2%; Score 47; DB 12; Length 228;

Best Local Similarity 70.0%; Pred. No. 2.2;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 CSDRRRCRXC 12

I: |||||:|

Db 66 CLDERCRRRC 75

RESULT 4

Q98204

ID Q98204

AC Q98204

DT 01-FEB-1997 (TREMBlrel. 02, Created)

DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)

DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)

DE MC036R.

GN MC036R.

OS Molluscum contagiosum virus subtype 1 (MCV1).

OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

OC Molluscipoxvirus.

OX NCBI_TaxID=10280;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96325459; PubMed=8670425;

RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,

RA Moss B.;

RT "Genome sequence of a human tumorigenic poxvirus: Prediction of

RT specific host response-evasion genes.";

RL Science 273:813-816(1996).

RN [2]

RP SEQUENCE FROM N.A.

RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,

RA Moss B.;

RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; U60315; AAC55164.1; -.

SQ SEQUENCE 1057 AA; 117533 MW; 7B18A5B7BC769EB2 CRC64;

Query Match 66.2%; Score 47; DB 12; Length 1057;

Best Local Similarity 70.0%; Pred. No. 8;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 CSDRRRCRXC 12

I: |||||:|

Db 895 CLDERCRRRC 904

RESULT 5

Q92VC2

ID Q92VC2

AC Q92VC2

DT 01-DEC-2001 (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE HYPOTHETICAL PROTEIN SMB21282.

GN SMB21282.

OS Rhizobium meliloti (Sinorhizobium meliloti).

OC Plasmid pSymB (megaplasmid 2).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396508; PubMed=11481431;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA Vorhoeelter F.J., Hernandez-Lucas I., Becker A., Cowie J.,
RA Golding B., Puhler A.;
RT "The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-
fixing endosymbiont Sinorhizobium meliloti.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR EMBL; AL603644; CAC49186.1; -
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 174 AA; 18995 MW; B242086334C12E14 CRC64;

Query Match 64.8%; Score 46; DB 16; Length 174;
Best Local Similarity 63.6%; Pred. No. 2.6;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCSDRRRCRX 12
Db 44 CCLRRRCRLPC 54

RESULT 6
Q69370 .PRELIMINARY; PRT; 399 AA.
AC Q69370;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE GLYCOPROTEIN GI.
OS Cercopithecine herpesvirus 2.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae.
OX NCBI_TaxID=10317;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93298054; PubMed=8390827;
RA Eberle R., Zhang M., Black D.;
RT "Gene mapping and sequence analysis of the unique short region of the
simian herpesvirus SA 8 genome.";
RL Arch. Virol. 130:391-411(1993).
DR EMBL; L05608; AAA46179.1; -
DR InterPro; IPR002874; Herpes_GI.
DR Pfam; PF01688; Herpes_GI; 1.
SQ SEQUENCE 399 AA; 42250 MW; D23B85CA9E454FF4 CRC64;

Query Match 64.8%; Score 46; DB 12; Length 399;
Best Local Similarity 63.6%; Pred. No. 5.2;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCSDRRRCRX 11
Db 306 SCCLRRRCRRR 316

RESULT 7
Q9HAU7 .PRELIMINARY; PRT; 488 AA.
AC Q9HAU7;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE SIALIC ACID-SPECIFIC 9-O-ACETYLESTERASE I.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Zhu H., Li J.M., Sha J.H.;
RT "A novel gene of sialic acid-specific 9-O-acetylesterase I from human
testis.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF300796; AAG15386.1; -
SQ SEQUENCE 488 AA; 54572 MW; DB030C82DA44916 CRC64;

Query Match 64.8%; Score 46; DB 4; Length 488;
Best Local Similarity 60.0%; Pred. No. 6.1;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACCSDRRRCRX 10
Db 407 SCCSDHRCKW 416

RESULT 8
Q9NT71 .PRELIMINARY; PRT; 521 AA.
AC Q9NT71;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE HYPOTHETICAL 58.1 KDA PROTEIN (FRAGMENT).
GN DKEZP761A051.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC OTTENWAEGLER B., Obermaier B., Gassenhuber J., Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL137496; CAB70771.1; -
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 521 AA; 58084 MW; B732FCBAFD82FED8 CRC64;

Query Match 64.8%; Score 46; DB 4; Length 521;
Best Local Similarity 60.0%; Pred. No. 6.4;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACCSDRRRCRX 10
Db 440 SCCSDHRCKW 449

RESULT 9
Q9HAT2 .PRELIMINARY; PRT; 523 AA.
AC Q9HAT2;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE SIALIC ACID-SPECIFIC ACETYLESTERASE II.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Zhu H., Zhou Z.M., Sha J.H.;
RT "A novel gene of human sialic acid-specific 9-O-acetylesterase from
testis.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF303378; AAG14897.1; -

SQ SEQUENCE 523 AA; 58315 MW; B72CF69636DBFED8 CRC64;

Query Match 64.8%; Score 46; DB 4; Length 523;
 Best Local Similarity 60.0%; Pred. No. 6.4;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACCSDRRCRX 10
 :||||| :
 DB 442 SCCSDRHCKW 451

RESULT 10
 P70665 PRELIMINARY; PRT; 541 AA.
 AC P70665; Q61044;
 DT 01-FEB-1997 (TReMBLrel. 02, Created)
 DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE SIALATE O-ACETYLESTERASE PRECURSOR (EC 3.1.1.53) (SIALIC ACID-SPECIFIC 9-O-ACETYLESTERASE) (YOLK SAC PROTEIN 2).
 DE YSG2
 Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (LYSOSOMAL ISOFORM).
 RC STRAIN-DBA/2 X C57BL/6;
 RX MEDLINE=97078679; PubMed=8918804;
 RA Stoddart A., Zhang Y., Paige C.J.;
 RA "Molecular cloning of the cDNA encoding a murine sialic acid-specific 9-O-acetyltransferase and RNA expression in cells of hematopoietic and non-hematopoietic origin."
 RT 9-O-acetyltransferase and RNA expression in cells of hematopoietic and non-hematopoietic origin.
 RL Nucleic Acids Res. 24:4003-4008(1996).
 RN [2]
 RP SEQUENCE FROM N.A. (LYSOSOMAL ISOFORM).
 RX MEDLINE=96278803; PubMed=8662838;
 RA Guimarães M.J., Bazar J.F., Castagnola J., Diaz S., Copeland N.G., Gilbert D.J., Jenkins N.A., Varki A., Zlotnik A.;
 RA "Molecular cloning and characterization of lysosomal sialic acid O-acetyltransferase."
 RT acetyltransferase.
 RL J. Biol. Chem. 271:13697-13705(1996).
 RN [3]
 RP SEQUENCE FROM N.A. (CYTOSOLIC ISOFORM).
 RC TISSUE-B-CELL;
 RX MEDLINE=99395134; PubMed=10464298;
 RA Takenatsu H., Diaz S., Stoddart A., Zhang Y., Varki A.;
 RA "Lysosomal and cytosolic sialic acid 9-O-acetyltransferase activities can be encoded by one gene via differential usage of a signal peptide-encoding exon at the N terminus."
 RT J. Biol. Chem. 274:25623-25631(1999).
 CC -1- FUNCTION: CATALYZES THE REMOVAL OF O-ACETYL ESTER GROUPS FROM POSITION 9 OF THE PARENT SIALIC ACID, N-ACETYLNEURAMINIC ACID.
 CC -1- CATALYTIC ACTIVITY: N-ACETYL-O-ACETYLNEURAMINATE + H(2)O = N-ACETYLNEURAMINATE + ACETATE.
 CC -1- ENZYME REGULATION: INHIBITED BY DIISOPROPYL FLUOROPHOSPHATE AND DIETHYL-P-NITROPHENYL PHOSPHATE (BY SIMILARITY).
 CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A SMALL SUBUNIT AND A LARGE SUBUNIT (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: LYSOSOMAL AND CYTOPLASMIC.
 CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS; LYSOSOMAL ISOFORM (SHOWN HERE) AND CYTOSOLIC ISOFORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: THE LYSOSOMAL ISOFORM IS WIDELY EXPRESSED. THE CYTOSOLIC ISOFORM SHOWS A MORE RESTRICTED DISTRIBUTION WITH HIGHEST EXPRESSION IN BRAIN AND OVARY AND LOWER LEVELS IN LIVER AND THYMUS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN MATURE B-CELL LINES BUT NOT IN LESS MATURE LINES.
 CC -1- PFM: THE TWO SUBUNITS ARE DERIVED FROM A SINGLE PRECURSOR BY PROTEOLYTIC CLEAVAGE (BY SIMILARITY).
 CC -1- PFM: THE LYSOSOMAL ISOFORM IS GLYCOSYLATED.
 DR EMBL; U61183; AAC52880.1; -.

DR EMBL; X98625; CAA67214.1; -.
 DR EMBL; U40408; AAB07813.1; -.
 DR EMBL; AF156856; AAD55976.1; -.
 DR MGD; MGI:104803; Ysg2.
 KW Hydrolase; Serine esterase; Glycoprotein; Lysosome; Signal;
 KW Alternative splicing.
 FT SIGNAL 1 23 BY SIMILARITY.
 FT CHAIN 24 275 SIALATE O-ACETYLESTERASE SMALL SUBUNIT.
 FT CHAIN 276 541 SIALATE O-ACETYLESTERASE LARGE SUBUNIT.
 FT CARBOHYD 107 107 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 138 138 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 188 188 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 427 427 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 462 462 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARSPLIC 1 98 MISSING (IN CYTOSOLIC ISOFORM).
 FT CONFLICT 160 160 I -> T (IN REF. 2).
 SQ SEQUENCE 541 AA; 60775 MW; 944936C45C4A2E6B CRC64;

Query Match 64.8%; Score 46; DB 11; Length 541;
 Best Local Similarity 60.0%; Pred. No. 6.6;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACCSDRRCRX 10
 :||||| :
 DB 468 SCCSDRHCKW 477

RESULT 11
 Q922L0 PRELIMINARY; PRT; 541 AA.
 AC Q922L0; DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE YOLK SAC GENE 2.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC007136; AAH07136.1; -.
 SQ SEQUENCE 541 AA; 60779 MW; 17967AC9C7D03917 CRC64;

Query Match 64.8%; Score 46; DB 11; Length 541;
 Best Local Similarity 60.0%; Pred. No. 6.6;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACCSDRRCRX 10
 :||||| :
 DB 468 SCCSDRHCKW 477

RESULT 12
 Q99552 PRELIMINARY; PRT; 3570 AA.
 AC Q99552; DT 01-MAY-1997 (TReMBLrel. 03, Created)
 DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE MUCIN MUC5B (FRAGMENT).
 GN MUC5B.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;

RN RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RX MEDLINE=97166151; PubMed=9013550;
 RA Desseyn J.L., Guyonnet-Duperat V., Porchet N., Aubert J.P., Laine A.;
 RT "Human mucin gene MUC5B, the 10.7 kb large-central exon encodes
 RT various alternate subdomains resulting in a super-repeat. Structural
 RL evidence for a 11p15.5 gene family.";
 RL J. Biol. Chem. 272:3168-3178(1997).
 DR EMBL: 272496; CAA96577.1;
 FT NON_TER 1 1
 FT NON_TER 3570 3570
 SQ SEQUENCE 3570 AA; 361019 MW; DE04E4D727579312 CRC64;

Query Match 64.8%; Score 46; DB 4; Length 3570;
 Best Local Similarity 70.0%; Pred. No. 32;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 2 CCSDRRRCRXR 11
 DB 273 CCSDDRRCGR 282

RESULT 13
 O39261 PRELIMINARY; PRT; 496 AA.
 AC O39261;
 DT 01-JAN-1998 (TRENBLREL. 05, Created)
 DT 01-JAN-1998 (TRENBLREL. 05, Last sequence update)
 DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
 DE COUNTERPART OF HSV-1 GENE UL41 AND VZV GENE 17.
 GN 19.
 OS Equine herpesvirus 4.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicelloviruses.
 OX NCBI_TaxID=10331;
 RN [1]
 RP SEQUENCE OF 456-496 FROM N.A.
 RC STRAIN=NS80567;
 RX MEDLINE=94266155; PubMed=8206376;
 RA Riggo M.P., Onions D.E.;
 RT "Sequences of the ribonucleotide reductase-encoding genes of equine
 RT herpesvirus 4.";
 RL Gene 143:217-222(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NS80567;
 RX MEDLINE=98264497; PubMed=9603335;
 RA Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;
 RT "The DNA sequence of equine herpesvirus-4.";
 RL J. Gen. Virol. 79:1197-1203(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NS80567;
 RA Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;
 RL Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF030027; AAC59533.1;
 DR InterPro: IPR000513; Exo_N.I.
 DR InterPro: IPR002927; Virus_HS.
 DR Pfam: PF01550; Virus_HS.1
 SQ SEQUENCE 496 AA; 56312 MW; 0A992AD765905862 CRC64;

Query Match 63.4%; Score 45; DB 12; Length 496;
 Best Local Similarity 72.7%; Pred. No. 8.9;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ACCSDRRRCRXR 11
 DB 343 ACLSDRCGRYR 353

RESULT 14
 Q9BYU7 PRELIMINARY; PRT; 106 AA.
 AC Q9BYU7;
 DT 01-JUN-2001 (TRENBLREL. 17, Created)
 DT 01-JUN-2001 (TRENBLREL. 17, Last sequence update)
 DT 01-OCT-2001 (TRENBLREL. 18, Last annotation update)
 DE KERATIN ASSOCIATED PROTEIN.
 GN KRTAP4.13.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SCALP;
 RA Rogers M.A., Langbein L., Winter H., Ehmann C., Korn B., Schweizer J.;
 RT "Characterization of a cluster of human high/ultrahigh keratin
 RT associated proteins on chromosome 17q12-21.";
 RL Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ296168; CAC27563.1;
 DR InterPro: IPR002494; Keratin_B2.
 DR InterPro: IPR001368; TNFR_c6.
 DR Pfam: PF01500; Keratin_B2.1.
 DR PROSITE: PS00652; TNFR_NGFR.1; UNKNOWN.1.
 SQ SEQUENCE 106 AA; 11570 MW; 15CF43A4C80C4375 CRC64;

Query Match 62.0%; Score 44; DB 4; Length 106;
 Best Local Similarity 54.5%; Pred. No. 3.6;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 2 CCSDRRRCRXR 12
 DB 75 CCQTRCRTC 85

RESULT 15
 Q9BYQ5 PRELIMINARY; PRT; 193 AA.
 AC Q9BYQ5;
 DT 01-JUN-2001 (TRENBLREL. 17, Created)
 DT 01-JUN-2001 (TRENBLREL. 17, Last sequence update)
 DT 01-OCT-2001 (TRENBLREL. 18, Last annotation update)
 DE KERATIN ASSOCIATED PROTEIN 4.15 (FRAGMENT).
 GN KRTAP4.15.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SCALP;
 RA Rogers M.A., Langbein L., Winter H., Ehmann C., Korn B., Schweizer J.;
 RT "Characterization of a cluster of human high/ultrahigh keratin
 RT associated proteins on chromosome 17q12-21.";
 RL Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ406945; CAC27584.1;
 DR InterPro: IPR002494; Keratin_B2.
 DR InterPro: IPR001007; VWFC.
 DR Pfam: PF01500; Keratin_B2.1.
 DR PROSITE: PS01208; VWFC; UNKNOWN.1.
 FT NON_TER 1
 SQ SEQUENCE 193 AA; 20665 MW; CC6F157C824C90E5 CRC64;

Query Match 62.0%; Score 44; DB 4; Length 193;
 Best Local Similarity 50.0%; Pred. No. 5.9;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 1 ACCSDRRRCRXR 12
 DB 106 SCCVSRCCRSQC 117

Search completed: July 1, 2002, 12:26:40
Job time: 182 sec